PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES

Volume 14 August 15, 1928 Number 8

THE THERMAL DECOMPOSITION OF AMMONIA UPON MIXED SURFACES OF TUNGSTEN AND PLATINUM

By Robert E. Burk*

Department of Chemistry, Cornell University

Communicated July 5, 1928

It is well recognized[1] that the lowering of the heat of activation is a probable way in which contact catalysts can accelerate reactions. The writer[2] has recently discussed two definite ways in which the catalyst could do this: (1) by distorting the atoms with a resulting weakening of the relevant bond, (2) by bodily separating the atoms forming the bond.

In the case of the thermal decomposition of ammonia on tungsten, the lowering of the heat of activation seems to be the only rôle played by the catalyst.[3] This reaction offers, therefore, a possible means of studying the mechanisms (1) and (2) and, therefore, of estimating their relative importance.

Some previous experiments[2] designed to investigate the contribution of mechanism (1), where ammonia was decomposed in strong electric fields, could not be carried through with complete satisfaction, because for fields greater than about 150,000 volts per cm., the fine wires (necessary in order to get the strong fields) were fused by ionic bombardment. However, the present experiments, designed to bring out the second mechanism, have met with more success.

If the bonds of ammonia are weakened by stretching, particularly active reaction centers should be produced by a solid whose surface contains a mixture of atoms, some of which adsorb the hydrogen atoms of ammonia very strongly, and others adsorb the nitrogen atom very strongly, provided, of course, that their spacing is suitable. If some factor other than the heat of activation does not operate, such a surface would exhibit "promoter action," i.e., its activity would be greater than an equal area of surface of any of the constituents. If, on the other hand, the first mechanism (atomic distortion) were the only effect of the catalyst, one would not be justified in expecting promoter action with any specific mixture of catalysts, and the promoter could not be chosen "theoretically" in the present state of knowledge.

On the basis of these considerations a mixed surface of tungsten and platinum was chosen as the prospective "promoted" catalyst, previous work[4] having given the kinetic characteristics of the decomposition of ammonia upon the separate metals, and together with general knowledge of the chemical and adsorptive properties of the metals, having rendered it plausible that tungsten adsorbs nitrogen more strongly than hydrogen, and that the reverse is true for platinum.

"Surface alloys" were produced by evaporating one metal from a spiral on to a filament of the other metal mounted axially within it.[5]

Not only did the mixed surface cause a greater velocity of decomposition than the same area of either platinum or tungsten alone, over the whole range of temperature investigated, but the temperature coefficient was decidedly smaller for the mixed surfaces, in complete agreement with expectations on the basis of mechanism (2).

CATALYST	HEAT OF ACTIVATION (CALORIES/GM. MOL.)
Tungsten (wire *A*)	47,000
Platinum	130,000
Tungsten-platinum (wire *X*)	34,600
Tungsten (wire *C*)	48,000
Tungsten-platinum (wire *C*)	36,000

While the temperature measurements were inaccurate due to a small progressive alteration in the standard resistance of the wire, no possible method of calculating the results could obscure them, as was found by trial. In the case of wire *C*, the temperature coefficient was measured on the pure, well-degassed tungsten wire; the platinum was then put on the identical wire, which was never out of the vacuum after the preceding ammonia decomposition, and was not heated at any temperature except that used in the decomposition of ammonia. There was thus little opportunity for the structure of the underlying metal to change, and the possible spacing factor to vary. Both synthetic ammonia and that made from Kahlbaum's pure ammonium chloride and pure calcium oxide, using all glass apparatus, gave the same results.

* NATIONAL RESEARCH FELLOW.

[1] See, e.g., Hinshelwood, *Kinetics of Chemical Change in Gaseous Systems*, **1926**. Rideal and Taylor, *Catalysis in Theory and Practice*, p. 46 (second edition).

[2] These PROCEEDINGS, **13**, 719 (1927); *J. Phys. Chem.*, **30**, 1134 (1926).

[3] The argument for this will be developed in the forthcoming "Sixth Report of the Committee on Contact Catalysis," and also in the more complete account of these experiments, which is to be published in the *Journal of Physical Chemistry*.

[4] Hinshelwood and Burk, *J. Chem. Soc.*, **127**, 1105 (1925).

[5] Full experimental details will be published. For the general procedure apart from the evaporation see Burk, these PROCEEDINGS, **13**, 67 (1927).

603

THE CRYSTAL STRUCTURE OF TOPAZ

BY LINUS PAULING

GATES CHEMICAL LABORATORY, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated July 11, 1928

An investigation of the orthorhombic crystal topaz, $Al_2SiO_4F_2$, with x-rays has been reported by Leonhardt.[1] With the aid of data from Laue and rotation photographs he found for the edges of the unit of structure, containing $4Al_2SiO_4F_2$, the values $d_{100} = 4.64$ Å, $d_{010} = 8.78$ Å, and $d_{001} = 8.37$ Å, and determined the space-group of the crystal to be V_h^{16}. A deduction of the atomic arrangement, dependent on about fifteen parameters, was not attempted; for such complicated structures cannot be treated by the usual methods.

By making use of the coördination theory of ionic crystals, I have predicted a structure for topaz which agrees satisfactorily with the experimental data. This structure is described in the following paragraphs.

The coördination theory, which has been applied successfully[2] to brookite, the orthorhombic form of titanium dioxide, is based on the assumption that the anions in a crystal are constrained to assume positions about the cations such that they indicate the corners of polyhedra of which the cations mark the centers. These polyhedra are the fundamentally important constituents of the crystal; retaining their form essentially, they are combined by sharing corners, edges, and faces in such a way as to build up a crystal with the correct stoichiometrical composition. Thus in rutile, anatase, and brookite there occur octahedra of oxygen ions about titanium ions; in rutile each octahedron shares two edges with adjoining octahedra, in anatase four, and in brookite three.

The fundamental polyhedra for topaz were assumed to be an octahedron of anions (oxygen and fluorine) about each aluminum ion and a tetrahedron of oxygen ions about each silicon ion. The anion-anion distance determining the length of a polyhedron edge was taken to be 2.72 Å throughout.[3] By piling these polyhedra together one structure was found, the unit of which approximated that observed for topaz. This structure is shown in figure 1, in which four layers of polyhedra to be superimposed are represented. On considering the distribution of microscopic symmetry elements for this predicted atomic arrangement its space-group was found to be V_h^{16}. The directions of the three axes are fixed for this space-group; using Niggli's convention (used by Leonhardt also), the c-axis is normal to reflection planes, the a-axis normal to glide planes with translational component $b/2$, and the b-axis normal to glide planes with translational component $a/2 + c/2$. The edges of the unit of structure are accordingly determined: their values, assuming the octahedra and

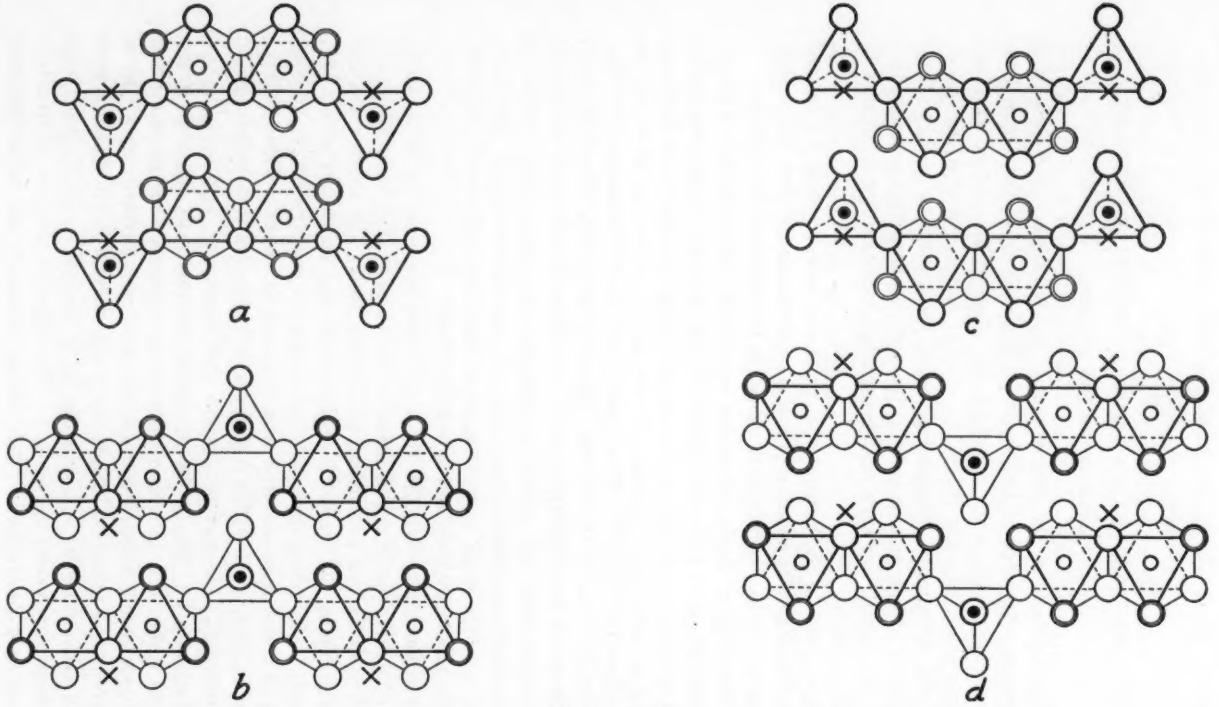


FIGURE 1

The arrangement of atoms in the crystal topaz. Large circles represent oxygen atoms, doubled large circles fluorine atoms, small open circles aluminum atoms, and small solid circles silicon atoms. The aluminum atoms are at the centers of the indicated octahedra, and the silicon atoms at the centers of tetrahedra. The four layers shown are to be superposed in the order *a*, *b*, *c*, *d*, with *d* uppermost. Crosses mark the projections of the corners of the unit of structure on the plane of the paper.

tetrahedra to be regular, are $d_{100} = 4.72$ Å, $d_{010} = 8.88$ Å and $d_{001} = 8.16$ Å. The unit contains $4Al_2SiO_4F_2$. The predicted structure leads to strong sixth order reflections from {100} and {001}, and strong fourth and eighth order reflections from {010}, in agreement with Leonhardt's observations.

The concordance of the predicted structure and the experimental data in regard to space-group symmetry, size of unit, and intensity of reflections from the pinacoids makes it highly probable that the correct structure of topaz has been found. The effect of deformation (using the methods applied to brookite) will next be considered, leading to the more accurate determination of the atomic arrangement. It is anticipated that this consideration will result in predicted values for the edges of the unit of structure in still better agreement with the observed ones (as with brookite), and that the predicted atomic arrangement will account completely for the observed intensities of reflection of x-rays from various crystallographic planes.

The atomic positions and the parameter values for the approximate structure with undistorted polyhedra are:

$8a$: $xyz;\ x\,y\,\frac{1}{2}-z;\ \frac{1}{2}-x\,\frac{1}{2}+y\,z;\ \frac{1}{2}+x\,\frac{1}{2}-y\,\frac{1}{2}+z;\ \bar{x}\bar{y}\bar{z};\ \bar{x}\,\bar{y}\,\frac{1}{2}+z;$
$\frac{1}{2}+x\,\frac{1}{2}-y\,\bar{z};\ \frac{1}{2}-x\,\frac{1}{2}+y\,\frac{1}{2}-z.$

$4c$: $u\,v\,\frac{1}{4};\ \bar{u}\,\bar{v}\,\frac{3}{4};\ \frac{1}{2}-u\,\frac{1}{2}+v\,\frac{1}{4};\ \frac{1}{2}+u\,\frac{1}{2}-v\,\frac{3}{4}.$

4Si in $4c$ with $u = \frac{11}{12}$, $v = \frac{1}{16}$;
8Al in $8a$ with $x = \frac{1}{12}$, $y = \frac{3}{8}$, $z = \frac{1}{12}$;
8F in $8a$ with $x = \frac{1}{12}$, $y = \frac{3}{4}$, $z = \frac{1}{12}$;
8O in $8a$ with $x = \frac{3}{4}$, $y = 0$, $z = \frac{1}{12}$;
4O in $4c$ with $u = \frac{1}{4}$, $v = 0$;
4O in $4c$ with $u = \frac{11}{12}$, $v = \frac{1}{4}$.

Each aluminum octahedron shares two edges with adjoining octahedra, and the four corners (oxygen ions) defining these edges are shared with tetrahedra. The other two corners, occupied by fluorine ions, are shared with octahedra. There are present in the structure SiO_4 groups; that is, tetrahedra having no elements in common with other tetrahedra; these tetrahedra share corners alone with octahedra. The oxygen ions, although crystallographically of three kinds, are from the standpoint of the coördination theory alike, for each is attached to two aluminum octahedra and one silicon tetrahedron. Each fluorine ion is shared between two octahedra.

The arrangement of the oxygen and fluorine ions is that of double hexagonal close-packing, so that Professor W. L. Bragg's methods of close-packing the large anions[4] could be applied to this crystal. But there are very many ways of distributing the cations with this anion arrangement, and the decision among them would have to be made by the laborious comparison of observed and predicted intensities of reflection, possibly involving the consideration of distortion of the anions from the close-

packed positions. Furthermore (as for brookite also), both hexagonal and double hexagonal close-packing are roughly compatible with the observed unit, increasing the number of possibilities to be considered. Bragg and West have, indeed, suggested with some reserve[4] that the anions in topaz are close-packed; their proposed type of close-packing is, however, incorrect, for it involves the identification of the hexagonal axis of the close-packed structure with the *a*-axis rather than the *b*-axis of topaz, and is, moreover, simple rather than double hexagonal close-packing.[5]

The further discussion of the structure of topaz will be published in the *Zeitschrift für Kristallographie.*

[1] J. Leonhardt, *Zeit. f. Krist.*, **59**, 216 (1924).

[2] Linus Pauling and J. H. Sturdivant, *Ibid.*, in press.

[3] Linus Pauling, *Jour. Amer. Chem. Soc.*, **49**, 763 (1927).

[4] W. L. Bragg and J. West, *Proc. Roy. Soc.*, **A114**, 450 (1927).

[5] Bragg and West report the dimensions of the unit of structure of topaz from their measurements to be $d_{100} = 4.64$ Å, $d_{010} = 8.79$ Å, and $d_{001} = 8.38$ Å, in good agreement with Leonhardt's values.

THE INTENSIVE DRYING OF LIQUIDS

BY SAM LENHER* AND FARRINGTON DANIELS

DEPARTMENT OF CHEMISTRY, UNIVERSITY OF CALIFORNIA

Communicated July 2, 1928

H. B. Baker has shown that liquids, gases and solids undergo a remarkable change in properties on prolonged and intensive drying with phosphorus pentoxide.[1] Experiments are described here which have been made with the purpose of repeating and verifying these results.

Tubes containing phosphorus pentoxide and organic liquids were sealed up and allowed to stand for a period of years. It was proposed to open the tubes after a drying time of 9 years (the time taken in Baker's experiments) and determine the physical properties of the liquid, such as boiling point, molecular weight, the capillary rise method, refractive index, etc. At the end of four and a half years it was felt that several tubes should be opened in order that the speed of drying by the Baker method might be tested. The results are reported below.

Drying was carried out in hard glass tubes having a wall 2 mm. thick, and an external diameter of 18 mm. The tubing chosen was free from visible imperfections, streaks, whorls, or cloudy areas which would indicate the existence of capillaries or fissures in the glass which might contain water. To make certain that the glass was free from capillaries it was heated to

redness (softening) along its entire length and the ends were drawn out. This precaution has since been described by Smits, de Liefde, Swart, and Claassen.[2] Phosphorus pentoxide was distilled in a current of dry oxygen using a platinized asbestos catalyst. The finely divided product gave no precipitate on standing for hours with silver nitrate solution showing that the resublimed pentoxide was free from traces of the lower oxides of phosphorus. The liquids, benzene and carbon tetrachloride, were carefully purified according to standard methods. The boiling points and freezing points checked with the accepted values.

The drying tubes were prepared as follows: A glass tube which had been fire heated to softening along its entire length was drawn off at one end and blown round. The liquid which had been standing for some weeks in contact with phosphorus pentoxide was introduced and the resublimed phosphorus pentoxide added. The pentoxide was introduced by plunging a dry glass tube into the redistilled product and transferring it to the tube of liquid where it was forced out of the tube with a glass rod. In this way the pentoxide did not come in contact with atmospheric moisture except possibly at the ends of the tube. The tube of liquid and drying agent was now sealed off, care being taken to thicken the glass where the final seal was made. The tubes containing benzene were cooled down to freezing before the final sealing off to prevent decomposition or explosion when the tube was sealed off. The drying was carried out at room temperature (18–20°C.). About 2–3 grams of phosphorus pentoxide were used in each tube.

Although it was proposed to measure a number of physical properties of these intensively dried liquids, the first test, the determination of the boiling point, showed that no additional measurements were necessary at this time. The boiling point was measured in two ways. To reproduce Baker's conditions, the tube was broken open at one end and put in a bath of sulphuric acid. The temperature of the bath was slowly raised and the temperature of the liquid and vapor noted when ebullition commenced. The second apparatus was a modified boiling-point apparatus of Mathews.[3] It consisted of a glass tube 1 cm. in diameter having a constricted lower end into which was fused a spiral heating coil. The platinum wire heating element was operated on 6 volts. About 5 cc. of the dried liquid was decanted from the drying tube into the dried boiling-point apparatus for a determination.

In every case the drying tube was opened directly to the air, as Baker has shown[4] that in the case of intensively dried benzene and carbon tetrachloride the recovery of the normal state is not rapid in comparison with the time of measurement (1 hour).

Tube 1. Benzene.—This tube of benzene and phosphorus pentoxide was prepared in the first week of December, 1923. It stood at room tempera-

ture until it was opened on May 16, 1928. Thus the drying time was $4^1/_2$ years.

The boiling point of the liquid in the drying tube was found to be 78.1°C. at 735.75 mm. or 80.67° at 760 mm. It was found that the dried liquid could be easily superheated 1.5°C. giving an apparent boiling point of as high as 82.12°C. at 760 mm., while the temperature of the vapor never rose above 78.2°C. at 735.75 mm. The boiling point of the intensively dried benzene measured in the platinum wire boiling-point apparatus was 78.1°C. at 735.75 mm. or 80.67° at 760 mm. The temperature of the liquid and vapor was constant, showing a variation of less than 0.03°C. The thermometers were standardized against B. S. thermometer No. 36497.

After the boiling-point determinations, the phosphorus pentoxide in the drying tube was treated with water in a calorimeter to determine how much unused pentoxide remained in the tube after the long period of drying. The pentoxide was still in a finely divided state dispersed in the benzene and no sign of the formation of metaphosphoric acid could be observed. The test showed that 1.6 grams of phosphorus pentoxide were still present. This value is probably 30% too low in view of the difficulty experienced in transferring all of the pentoxide from the drying tube to the calorimeter without heat loss.

Tube 2. Benzene.—This was a double-arm tube, originally designed for distilling the liquid from one arm to the other to speed up the drying process. It was prepared in June, 1924. As this tube was opened on May 16, 1928, the drying time was 3 years, 11 months.

The boiling point determined in the sulphuric acid bath was 78.1°C. at 735.60 mm. or 80.69° at 760 mm. The benzene could be superheated 1.3°C. giving a boiling point as high as 82.03° at 760 mm. The boiling point in the platinum wire apparatus was 78.1°C. at 735.60 mm. or 80.69 at 760. There was less than 0.03°C. difference between the liquid and the vapor 2 cm. above it.

Tube 3. Carbon tetrachloride.—Tube 3 was prepared in November, 1923. In June, 1924, the tube was broken open and the normal boiling point of carbon tetrachloride was found. The liquid was at once sealed up again in a new tube (dried and heated) with a fresh portion of redistilled and oxidized phosphorus pentoxide. When the tube was opened on May 16, 1928, the drying time was over $4^1/_2$ years, with a continuous drying time of 3 years, 11 months.

The boiling point with the sulphuric acid bath was 74.3°C. at 735.70 mm. or 76.75 at 760 mm. The carbon tetrachloride could be easily superheated 1.5°C. before ebullition began. Using the platinum wire apparatus a boiling point of 74.3 ± 0.1°C. was found in the liquid phase, and 74.3 ± 0.2°C. (at 735.70 mm.) in the vapor phase.

A calorimetric determination of the amount of phosphorus pentoxide remaining in the drying tube showed that 3.8 grams of pentoxide were unused. This quantity of phosphorus pentoxide is probably 30% too low.

Summary.—In three experiments described here no abnormal rise in the boiling points of benzene and carbon tetrachloride was obtained after four to four and a half years of intensive drying. The organic liquids were sealed off with phosphorus pentoxide in glass tubes and kept at room temperature.

It may be remarked that there remain a number of tubes of glass and quartz containing organic liquids sealed up with phosphorus pentoxide and it is proposed to open them several years later when they are at least as old as the tubes described by Baker.

The work was carried out in the Laboratory of Physical Chemistry, University of Wisconsin.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] H. B. Baker and M. Baker, *J. Chem. Soc.*, **101**, 2339 (1912); H. B. Baker, *J. Chem. Soc.*, **121**, 568 (1922); A. Smits, *J. Chem. Soc.*, (1924–1926).

[2] *J. Chem. Soc.*, **129**, 2657 (1926).

[3] *Trans. Amer. Electrochem. Soc.*, **19**, 81 (1911).

[4] *J. Chem. Soc.*, **1922**, loc. cit.

A PROGRESSION RELATION IN THE MOLECULAR SPECTRUM OF OXYGEN OCCURRING IN THE LIQUID AND IN THE GAS AT HIGH PRESSURE

BY OLIVER R. WULF*

CHEMICAL LABORATORY, UNIVERSITY OF CALIFORNIA

Communicated June 23, 1928

During a study of the existing data on the absorption of light by oxygen gas and by liquid oxygen, the author has noticed a new progression relation in the near ultra-violet. In the gas at moderate pressures there begins to appear absorption[1] which is quite distinct from the Schumann-Runge system and from the atmospheric bands.[2] This has been observed several times in the past,[1] and was found to occur also in liquid oxygen.[3] In an extensive investigation of the absorption of liquified gases in the ultra-violet, Ciechomski[3] obtained considerable data on liquid oxygen. The more recent work of Shaver,[2] who studied this absorption in the visible and ultra-violet and in both the gas and liquid, leaves no doubt that the liquid exhibits this absorption in practically the same manner as the gas, any differences being of a minor nature insofar as the general character of the absorption is concerned. There are a number of bands, appearing

just to the red of complete absorption (which begins in the vicinity of 2400 Å), for which Ciechomski has given the approximate positions. The author has noticed that a majority of these are situated in so regular a way that they undoubtedly constitute a progression of bands. The bands have not been resolved, but Ciechomski observed in those best defined that each band appeared to consist of a group of three narrower bands. The work of Shaver has confirmed this. With insufficient definition the appearance is that of a single rather broad band.

The origin of these bands is not known, but the author has recently given reasons[4] for believing that the new absorption, represented by scattered bands entering in the gas at high pressures and present in the liquid, is due, not to O_2 but to a polymer of O_2, the molecule O_4.[5]

The above-mentioned bands which occur on the edge of the complete absorption show a strong resemblance to the type of which the bromine bands, and the Schumann bands of O_2, are illustrations. Their intensity increases *rapidly* to shorter wave-lengths, for any one length of light path and concentration of oxygen, only about four being discernible, the first very weak and the last very strong, followed by complete absorption. Ciechomski was able to extend his observations by using different lengths of path and using liquid air as well as liquid oxygen to obtain different oxygen concentrations. Their character suggests that these bands are due to an electron transition in which the molecular binding undergoes a great weakening and for which the most probable transition for absorption from the non-vibrating normal state is to one of the higher vibration levels of the excited state or into the continuous spectrum representing dissociation, in accordance with the theory of Franck.[6]

In the following table the first column contains Ciechomski's numbering of the bands and in the second column their long and short wave-length limits, these two columns being taken directly from Ciechomski's Table 2. The means of these limits are tabulated in the third column and the wave numbers of the means in the fourth. In the fifth column are contained the differences in wave numbers between those bands that have been numbered arbitrarily in the last column.

The bands numbered 11 to 20 by Ciechomski are the ones of particular interest here. Bands 8, 9 and 10 are quite apart from these, both in character and position. The same is true also of band 13, as an inspection of the reproductions of Ciechomski's plates shows. It is narrow and weak, and apparently different from its immediate neighbors. This is not surprising, since other scattered absorption is found. Ciechomski was able to observe the triplet-like structure in bands 12, 14, 15, 16 and 17.

The bands numbered by the author arbitrarily from 1 to 9 constitute the progression and they are plotted in the upper portion of the figure. Ciechomski was unable to measure the limits of the last two and these are

therefore probably somewhat less accurately placed. To use the mean of the limits, that is, the center of the broad bands as representing their position is, of course, arbitrary, but since their structure is not known, there does not seem to be a better procedure. As the figure shows, these nine bands fall fairly satisfactorily into a converging progression.

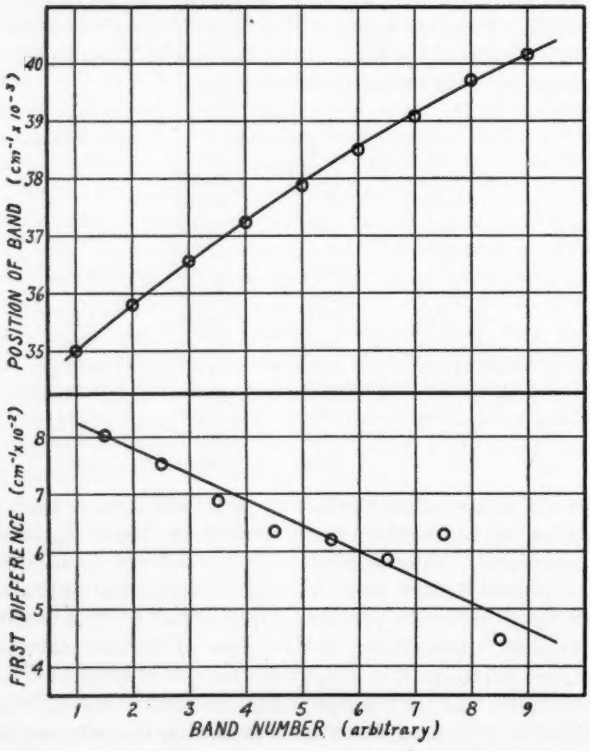
The first differences given in the table are plotted in the lower portion of the figure. These come rather far from showing the usual uniform decrease characteristic of a band progression; yet, considering that the bands are not resolved nor easy to measure, the fluctuations are perhaps not surprising. From the figure it seems possible that the wave number for band 8 is somewhat too high. If so, this would account partly for the large fluctuations in the last two first differences.

TABLE 1

NO. DATA OF CIECHOMSKI	BAND LIMITS, Å	MEAN POSITION Å	MEAN POSITION CM.$^{-1}$	FIRST DIFFERENCE CM.$^{-1}$	PROGRESSION NO.
8	3825–3800	3812.5	26230		
9	3625–3600	3612.5	27682		
10	3450–3430	3440	29070		
11	2875–2839	2857	35002		1
12	2809–2777	2793	35804	802	2
13	2767–2760	2763.5	36186		
14	2752–2719	2735.5	36556	752	3
15	2700–2670	2685	37244	688	4
16	2655–2625	2640	37879	635	5
17	2610–2585	2597.5	38499	620	6
18	2571–2546	2558.5	39085	586	7
19	2518	2518	39714	629	8
20	2490	2490	40161	447	9

There are a number of cases where the first differences in band progressions decrease approximately linearly toward the beginning of the continuous spectrum.[7] No very satisfactory attempt can be made to extrapolate the present data. However, the straight line inserted in the figure (the upper curve is a parabola corresponding to this) would appear to be a possible representation of the course of the first differences. If these progress approximately linearly to the axis of abscissae, the extrapolation indicates that the beginning of the continuous absorption is in the neighborhood of 2350 Å, at about the 19th band in this arbitrary numbering. An earlier, much less certain estimation of a different kind from three semi-quantitative data of Liveing and Dewar[2] indicated the vicinity of 2500–2400 Å. As has been mentioned, this absorption is believed to be due to the molecule O_4 and to correspond to the oscillation O—O_3. In accordance with our knowledge of the behavior of other molecules, the continuous spectrum probably represents the dissociation of the molecule in the act of absorption. If the assumption is made that the dissociation is into unexcited constituents, the position of the beginning of the con-

tinuous can be predicted, since it must correspond to the energy of dissociation of O_4 into O_3 and O. This can be calculated thermochemically from the heat of formation of ozone,[8] the heat of dissociation of O_4 into two O_2 molecules,[5] and the heat of dissociation of O_2.[7] The same result would apply nearly as well if the dissociation were into normal ozone molecule and excited oxygen atom, provided the excitation of the latter is limited to one of the states of the basic triple P level, as this would only make a difference of a few hundredths of a volt. If the excitation were higher,



the calculation would be in doubt so long as the excitation was unknown. Also, if the dissociation were into normal oxygen atom and excited ozone molecule, the calculation could not be made, since nothing is known about the electronic levels of the O_3 molecule. Carrying out the calculation one finds that, for the first two possibilities, the beginning of the continuous spectrum would be expected at approximately 2470 Å. There are uncertainties in all of these three estimations, but their rough agreement may be significant.

The rapid increase in intensity of the bands as they progress toward the continuous absorption makes it seem quite possible that the probability of transition to the continuous is so great that, even in moderate lengths of path in oxygen at moderate pressures, there may be appreciable intensity of absorption in the continuous region, for instance, 2300–2100 Å.

This progression of bands is far to the red of any that could be due in absorption to the Schumann-Runge system of the O_2 molecule at ordinary temperatures. Furthermore, the first differences for bands 1, 2 and 3 are larger than the largest first differences for the Schumann bands, and these three are probably not the first bands by any means of the new progression. The atmospheric bands of oxygen, in the red and yellow portion of the visible, have a much larger first difference initially. This system is not on as secure a foundation yet as is the Schumann-Runge system. Now it is interesting to note that the first difference of the atmospheric bands progression in the region of these new bands, estimated by making a long linear extrapolation of the former is not so greatly different from the new bands, thus at about 2850 Å being 886 cm.$^{-1}$ as compared with 800 cm.$^{-1}$. However, the intensity of the atmospheric bands *decreases* rapidly in the one progression known, in accord with the fact that the moment of inertia of the molecule in the excited state is only slightly different than in the normal state. Just the reverse is true in the case of the Schumann bands, and the new ones as mentioned above are of this latter type. There appears to be no reason for expecting any connection between the atmospheric bands and the new ones. The triplet-like structure of the new ones is, of course, still obscure and bears no resemblance to the bands of these other systems.

The author gratefully acknowledges his indebtedness to Professor Raymond T. Birge for his helpful criticism and suggestions.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] Janssen, *Compt. rend.*, **102**, 1352 (1886); *Ibid.*, **120**, 1306 (1895); Baceii, *N. Cim.*, **9**, 177 (1899); Liveing and Dewar, *Proc. Roy. Soc.*, **46**, 222 (1889); *Phil. Mag.*, **26**, 286 (1888); *Ibid.*, **34**, 205 (1892); *Ibid.*, **40**, 268 (1895); Shaver, *Proc. Roy. Soc. Canada*, **15**, sec. III, 7 (1921).

[2] Concerning the latter, see, for example, reference 7.

[3] Olzewski, *Wied. Ann.*, **42**, 663 (1891); Liveing and Dewar, loc. cit.; Ciechomski, Dissertation, Freiburg (Schweiz), **1910**. For an abstract, see *Beibl. Ann. Physik*, **39**, 240 (1915); Shaver, loc. cit.

[4] Wulf, *Proc. Nat. Acad. Sci.*, **14**, 356 (1928).

[5] Lewis, *Jour. Am. Chem. Soc.*, **46**, 2027 (1924).

[6] Franck, *Zs. für physik. Chemie*, **120**, 144 (1926).

[7] Birge and Sponer, *Phys. Rev.*, **28**, 259 (1926).

[8] Jahn, *Zs. anorg. Chem.*, **60**, 337 (1908); Kailan and Jahn, *Ibid.*, **68**, 243 (1910).

THE HEAT OF DISSOCIATION OF OXYGEN AS ESTIMATED FROM PHOTOCHEMICAL OZONIZATION

BY OLIVER R. WULF*

CHEMICAL LABORATORY, UNIVERSITY OF CALIFORNIA

Communicated June 23, 1928

The heat of dissociation of oxygen has been estimated by Warburg[1] and by Born and Gerlach,[2] using, in both cases, the data of Warburg[3] on the photochemical formation of ozone. The assumptions made, however, were different, and quite different results were obtained, 138,000 calories in the former case as compared with 162,000 calories in the latter.

The author believes that the estimation made by Born and Gerlach rests upon a mistaken assumption. The agreement of their result with the more recent determinations made by Hogness and Lunn[4] from positive-ray studies, and by Birge and Sponer[5] from spectroscopic data, appears to be fortuitous. The author further believes, from a recent study of the existing data on the absorption of light by gaseous and liquid oxygen,[6] that the estimation made by Warburg[1] pertains probably to the molecule O_4 rather than O_2. The purpose of this paper is to give briefly the evidence which supports these statements.

During his classical researches, which constitute the foundation of modern photochemistry, Warburg studied the formation of ozone from oxygen at wave-lengths 2530 Å and 2070 Å.[3] At the latter he found conditions which nearly satisfied the photochemical equivalence equation, while at the former this was not the case. Wave-length 2530 Å was the longest used, because, even at the high pressures worked with, oxygen does not absorb appreciably much further to the red.

In the estimation made by Born and Gerlach, one finds the important basis of the calculation to be contained in the statement that the longest wave-length which forms ozone is 2530 Å, making reference to the work of Warburg.[3] But there does not seem to be sufficient reason for believing that this is the case, at least not because of any energy considerations. The circumstance that oxygen does not absorb much to the red of this, even at the highest pressure used by Warburg, does not appear to permit taking this to be a long wave-length limit for ozone formation. Light of longer wave-length, i.e., corresponding to lower energy, might also yield ozone if the gas could absorb such radiation. It seems probable, that this would actually be found to be true at still-higher pressures. At a pressure of 400 kgm./sq. cm. λ 2820 Å, another source used by Warburg, is not appreciably absorbed.

The estimation of Born and Gerlach was to be taken as an upper limit because of the nature of the cyclic process used, there being two unknown

quantities contained in it. From the present point of view, if indeed the absorption of 2530 Å were by the O_2 molecule, the result of such a computation would be an upper limit in still another sense, namely, that there does not appear to be sufficient reason for believing that the quantum corresponding to 2530 Å is the smallest unit of energy, which will excite the absorbing molecule to ozone formation. Thus the estimate, viewed simply on its own merits, might be much too high.

Furthermore, this absorption of oxygen gas at high pressure in the near ultra-violet, seems to be due to a polymer of O_2,[6] the molecule O_4, as will be mentioned further. Wave-length 2530 Å lies close to the 8th (arbitrary) member of the band progression recently noticed by the author[7] during a study of the existing data on the absorption of oxygen. This wave-length lies thus simply in a portion of this *discontinuous* spectrum where, at these pressures, the intensity of absorption is still appreciable.

Warburg, on the other hand, concluded, from the agreement with the conditions of the photochemical equivalence equation at λ 2070 Å, that the primary photochemical act at this wave-length was the dissociation of the absorbing molecule. This principle has received important experimental support since that time, so that it seems very probable that, in endothermic photochemical reactions, the absorption of light, which results in a reaction of unit quantum efficiency, represents the dissociation of the molecule in the elementary photochemical act.

But so far as is known at present, molecules are dissociated in an elementary act by light absorption only when such absorption lies in the continuous spectrum that stretches to shorter wave-lengths from the point of confluence of a band progression.† The familiar ultra-violet absorption of oxygen, the Schumann-Runge band system, possesses no such continuous absorption in the region of 2070 Å at room temperature. Indeed, the electron transition on which this system is built is at 2026 Å (49,356 $cm.^{-1}$)[8], that is, this is the position of the 0—0 band. Thus all of the progression built on the normal non-vibrating molecule as the initial state lies beyond 2070 Å, the continuous absorption belonging to this, setting in first at 1751 Å. And the first members of this progression are relatively very weak, the absorption increasing as the progression advances to shorter wave-lengths. Absorption due to the normal molecule possessing one unit of vibrational energy will probably be present to some extent at room temperature but much weaker than the former. The second band of this progression, the 1—1 band, has its edge at 2062 Å, and could take some part in the absorption of this group of zinc lines around 2070 Å. It is difficult to estimate the part played in the absorption by the states above[9] the normal vibrationless state at this temperature and pressure. Offsetting the very small concentration in the upper states (even for the state of one unit of vibration the Boltzmann factor is approximately 5 ×

10^{-4}), there is the rapid increase in the probability of transition to the lower vibration states of the excited molecule with increase in vibrational energy of the initial state. From the available data it seems, however, that such absorption with the short light paths employed in the photochemical experiments, even at the high pressures used, will be relatively weak at room temperature. Furthermore, such absorption does not represent dissociation. It does not seem that it can account for the major part of the intense absorption of 2070 Å by oxygen gas under these conditions.

However, in oxygen gas at high pressures and in liquid oxygen a new absorption enters,[10] distinct from the Schumann-Runge system, and in which it is probable that a continuous spectrum begins in the vicinity of 2400 Å, at the point of confluence of a band progression, and stretches to shorter wave-lengths. Thus 2070 Å lies in this continuous spectrum. The author has discussed the structure of this absorption[7] and has given reasons for believing that it is due to the molecule O_4.

Thus 2070 Å appears, indeed, from this point of view to be an upper limit for a dissociation energy, but not of the molecule O_2, rather of the molecule O_4 into the components O_3 and O. The *beginning* of the continuous spectrum, under certain conditions, might be an actual measure[7] of this dissociation energy. A wave-length between 2070 Å and 2400 Å would also be expected to yield unit quantum efficiency and it would be interesting to know whether the wave-length 2220 Å, used by Warburg in other recent photochemical work, really would do so. It would be situated with respect to 2070 Å for oxygen in much the same way as 2530 Å is to 2070 Å in the case of HI, where for both wave-lengths, Warburg[11] found the conditions of the equivalence equation satisfied; and we know today that both these lie in the continuous spectrum which represents the dissociation of the molecule HI in the primary act of absorption.[12] The molecule HBr would serve as a similar illustration.

It should be pointed out, however, that, if this absorption, which the author believes is due to the molecule O_4, is due rather to an unknown level in the molecule O_2, Warburg's work points to a lower value for the heat of dissociation of the molecule than that now generally accepted.

There does not seem to be any way of obtaining information regarding the heat of dissociation of oxygen from photochemical ozonization, which is not also afforded by the interpretation of the spectral data alone.

* NATIONAL RESEARCH FELLOW IN CHEMISTRY.

[1] Warburg, *Z. f. Elektrochemie*, **26**, 58 (1920).
[2] Born and Gerlach, *Z. f. Physik*, **5**, 433 (1921).
[3] Warburg, *Berl. Ber.*, p. 872 (1914); see also *Z. f. Elektrochemie*, **7**, 133 (1921).
[4] Hogness and Lunn, *Phys. Rev.*, **27**, 732 (1926).
[5] Birge and Sponer, *Ibid.*, **28**, 259 (1926).
[6] Wulf, *Proc. Nat. Acad. Sci.*, **14**, 356 (1928).

[7] Wulf, *Ibid.*, **14**, 609 (1928).

[8] "Molecular Spectra in Gases," *Bull. Nat. Res. Council*, No. 57, p. 232.

[9] Füchtbauer and Holm, *Phys. Zeitschr.*, **26**, 345 (1925); von Wartenberg, *Ibid.*, **11**, 1168 (1910).

[10] For references to this see footnotes 1 and 3 in preceding article by author.

[11] Warburg, *Berl. Ber.*, p. 300 (1918).

[12] Coehn and Stuckardt, *Z. f. phys. Chem.*, **91**, 722 (1916); Bonhoeffer and Steiner, *Ibid.*, **122**, 287 (1926); Tingey and Gerke, *J. Am. Chem. Soc.*, **48**, 1838 (1926); Bernard Lewis, *J. Phys. Chem.*, **32**, 270 (1928); Bonhoeffer and Farkas, *Z. f. phys. Chem.*, **132**, 235 (1928).

† *Note added in proof*—See, however, the recent paper by Bonhoeffer and Farkas, *Z. f. phys. Chem.*, **134**, 337 (1928) for another interesting possibility.

THE CRYSTAL STRUCTURES OF THE ALKALINE EARTH METALS

BY G. L. CLARK, A. J. KING AND J. F. HYDE

DEPARTMENT OF CHEMISTRY, UNIVERSITY OF ILLINOIS

Communicated July 6, 1928

Prior to the investigation here reported the only one of the alkaline earth metals which had been satisfactorily subjected to x-ray crystal analysis was calcium. Hull[1] reported a face-centered cubic lattice with an edge length of 5.56 A.U. Most of the difficulty with strontium and barium seems to have been in a method of preparation of the pure metals. In the present work calcium, strontium and barium were prepared with purities exceeding 99.9% and subjected to analysis of crystalline structure by the Hull powder method of x-ray spectrography.

The results of Hull on calcium were fully substantiated both as to type of lattice and constant.

Barium was prepared by subliming the metal from a mixture of barium oxide and aluminum in a vacuum sublimation furnace, the product being handled in an atmosphere of argon to avoid oxidation.[2] The samples for x-ray analysis were prepared by first pressing the sublimed crystals into a cylindrical pellet in a steel die, and then rolling the pellet into a sheet using oil to protect the metal from oxidation during the process of rolling. The thin sheets were then cut into very narrow ribbons under oil and quickly sealed into thin pyrex tubes with sodium chloride which served as a standard for measurement of the diffraction pattern. Several excellent spectrograms with sharp lines over the whole length of the film and free from fogging were obtained.

Analysis of the spectrograms shows that barium crystallizes in the cubic system, the unit cell of which contains two atoms, and whose constant is

5.04 A.U. The intensities of the lines are in agreement with those calculated for the *body-centered* cubic lattice. The density calculated from the above data is 3.52 which compares well with the value 3.5 given in the International Critical Tables and 3.64 obtained from density determinations made with very pure barium wire.[3]

Specimens of pure strontium were prepared in analogous fashion, the utmost precaution being employed to prevent oxidation, though this metal is not so easily oxidized as is barium. Instead of obtaining the sharp spectra observed for barium, the films were all considerably fogged and showed at the most seven ill-defined diffraction lines. Repeated efforts were made to increase grain size by mechanical and heat treatment, and to chill the specimen in liquid air, but the spectra were not improved. Inasmuch as the Kα-rays of molybdenum (42) may excite the fluorescent Kα-rays of strontium (38), spectra were also obtained with copper Kα rays from a Hadding-Siegbahn tube, again without improvement.

Since the announcement of these results were made at the Saint Louis meeting of the American Chemical Society in April, a paper has appeared by Simon and Vohsen[4] presenting the results of crystal analyses of the alkali metals and of strontium (the authors state that they had been unable to prepare pure barium). They assign a face-centered cubic lattice to strontium, with the constant 6.03 A.U. They believe that an hexagonal modification may be possible at high temperatures, since optical examination indicates that neither calcium nor strontium crystallize in the regular system. The spectra obtained in the present work do not correspond to any cubic lattice. Of the seven measurable lines possibly two might be made to agree with the structure published by Simon and Vohsen. Inasmuch as strontium is a transition step between calcium (face-centered cubic) and barium (body-centered cubic) it is perhaps not strange that it should behave in anomalous fashion with a transition point between two or more modifications at room temperature. Impurities or oxide layers cannot be said to be the cause of poor diffraction as has recently been demonstrated for potassium, for the metal upon chemical analysis was found to be more than 99.9% pure. The work is being continued with renewed interest in the light of the new work of Simon and Vohsen. Upon the basis of the present results it may be predicted that radium should crystallize on the body-centered cubic lattice.

[1] *Physical Rev.*, **17**, 42 (1921).

[2] A. J. King, Doctorate Thesis, Syracuse University, 1927. Danner, *J. Am. Chem. Soc.*, **46**, 2382 (1924). Ruff and Hartman, *Zeit. anorg. Chem.*, **133**, 29 (1924). Biltz and Huttig, *Ibid.*, **114**, 241 (1920). Dafert and Miklanz, *Monatsch.*, **34**, 1685 (1913). Guntz and co-workers, cf. *Bull. soc. Chim.*, **35**, 709 (1924).

[3] King, loc. cit.

[4] *Zeit. physikal. Chem.*, **133**, 165 (1928).

REFLECTION AND REFRACTION OF ELECTRONS BY A CRYSTAL OF NICKEL

BY C. J. DAVISSON AND L. H. GERMER

BELL TELEPHONE LABORATORIES

Read before the Academy April 23, 1928

In the April number of these PROCEEDINGS[1] we reported certain preliminary results obtained in experiments in which a homogeneous beam of electrons was directed against a {111}-face of a nickel crystal at various angles of incidence, and in which observations were made on the intensity of scattering in the plane of incidence as a function of bombarding potential and direction. We had found that the incident beam of electrons is regularly but selectively reflected from the crystal face. At a given angle of incidence the reflected beam is observed whenever the speed of the incident electrons is comprised within any of certain ranges, and within each of these ranges the intensity of the beam is characterized by a sharply defined maximum. The phenomenon was interpreted as the wave mechanics analogue of the regular selective reflection of monochromatic x-rays from a crystal face.

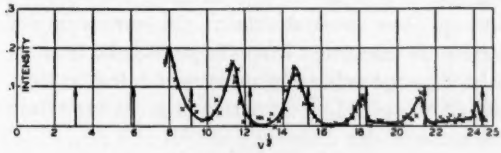


FIGURE 1

Variation of the intensity of the regularly reflected electron beam with bombarding potential, for 10° incidence—Intensity vs. $V^{1/2}$.

In the x-ray phenomenon the intensity of the reflected beam is a maximum when the wave-length of the incident beam satisfies the Bragg formula, when, that is, the wave-length has any one of the values, $\lambda = (2d/n) \cos \theta$, where d represents the distance between adjacent planes of atoms lying parallel to the surface of the crystal, θ the angle of incidence, and n any positive integer. A complete analogy between the phenomena of electron reflection and x-ray reflection would require that the Bragg formula should hold also in the case of electrons. This condition, however, is not satisfied; the wave-lengths at which the beam of reflected electrons attains its intensity maxima are not given by $\lambda = (2d/n) \cos \theta$.

This failure to conform to the Bragg law is illustrated in figure 1, which is figure 3 of our previous note (loc. cit.). Observations on the intensity of the reflected beam for angle of incidence 10 degrees are plotted in this figure

against the square-root of the bombarding potential which is proportional to the speed of bombardment, and therefore to the reciprocal of the wave-length. [$\lambda = h/mv = 12.2/V^{1/2}$ A, for electrons of moderate speed.] If the Bragg formula obtained, the maxima in this curve would occur at the values of $V^{1/2}$ given by

$$V^{1/2} = 12.2/\lambda = 12.2\left(\frac{n}{2d\cos\theta}\right)$$
$$= n \times 3.05, \text{ for } \theta = 10 \text{ deg.}, d = 2.03\ A.$$

These values are indicated by the arrows in the figure, and one notes a definite failure of the observed maxima to fall at the calculated positions.

A discrepancy of this sort was not unexpected. We had found in our first experiments[2] that electron diffraction beams do not coincide in position or in wave-length with their Laue beam analogues, and it was anticipated that the properties of the crystal responsible for these discrepancies would manifest themselves, in the case of electron reflection, in a departure from the Bragg law. On the other hand, it was expected from considerations of symmetry that the reflection would be regular, and it appears to be strictly so in all cases.

The single statement covering both reflection and diffraction is that for electrons of the speeds used in our experiments (bombarding potentials up to 600 volts) Bragg's law does not obtain; the wave-length of the beam of scattered electrons as calculated from the de Broglie formula is never the same (except in a special case to be mentioned later) as that of the corresponding beam of x-rays. It is a matter of great importance, of course, to discover the cause of this difference.

The suggestion was made by Eckart[3] on the basis of the results contained in our note to *Nature* that this failure of the Bragg law is to be ascribed to a refraction of the electron waves by the crystal. The same suggestion was made also by Bethe[4] who pointed out that such an effect could be readily accounted for on the wave theory of mechanics as a consequence of the change in potential energy experienced by the electrons on passing into the metal. This idea is a particularly attractive one as it is capable of accounting not only for wave-length and directional differences of the types observed, but is capable of accounting also for the well-established fact that the wave-length λ ($= h/mv$) of each of the diffraction beams satisfies the plane grating formula with respect to one or another of the atomic plane gratings to which the surface layer of atoms is equivalent.

Bethe calculated indices of refraction for nickel from the data contained in our note to *Nature*, and in our *Physical Review* article (loc. cit.) we made similar calculations and displayed the results in figure 18. Both of these sets of calculations were based upon the assumption that each electron beam is the analogue of the Laue beam of the same azimuth and order and

of the next *longer* wave-length, this being the correlation we then favored. These calculations yielded values of refractive index less than unity. We took occasion to point out that the correlation of x-ray and electron beams which we then favored was not necessarily the correct one, that if each electron beam were assumed to be the analogue of the adjacent Laue beam of *shorter* wave-length the indices of refraction would all be greater than unity, and that such values seemed inherently more acceptable. The discrimination was based upon the consideration that if the work function of the metal is equivalent to Φ volts one might reasonably expect the index of refraction of the metal for electrons to be given by $\mu = (1 + \Phi/V)^{1/2}$, where V represents the bombarding potential. Values of μ corresponding to this alternative correlation were not, however, given in our paper, nor were values of Φ given for either correlation.

A similar expression for the refractive index, to wit, $\mu = (1-V/E)^{1/2}$, was derived by Bethe (loc. cit.) from Schroedinger's equation. In this formula E represents the total energy of the electron (bombarding potential) and V the mean potential energy of the electron inside the crystal. The zero of potential is taken outside the crystal so that a negative value is expected for V and values greater than unity for μ.

In a more recent note Bethe[5] has calculated values of V corresponding to both correlations of electron and Laue beams and finds that the values corresponding to our alternative correlation show smaller departures from their mean value than those corresponding to the correlation which we originally favored. According to Bethe's calculations the most probable value of V is -15 volts. And this value he finds acceptable on the view that it represents not the Richardson work function of the metal, but the space average of potential within the crystal, a quantity which figures in the Fermi-Sommerfeld theory of thermionic emission.

To account for the absence of the analogues of certain of the principal Laue beams in this second or alternative correlation Bethe suggests, as we also suggested in our *Physical Review* article, that these beams suffer total reflection at the surface of the crystal and therefore fail to emerge; also that the intense beam occurring in the {100}-azimuth for bombarding potential 35 volts should be regarded as a diffraction beam rather than as a "grazing beam." We differed from Bethe in thinking that these considerations were inadequate to account for all of the missing beams.

The observations on the electron reflection beams reported in this and in our previous note were undertaken to determine which, if either, of the proposed correlations is the correct one. We hoped also to determine whether or not refraction in the ordinary optical sense is actually a property of the crystal. For in spite of the theoretical considerations favoring this view, the experimental evidence supporting it was very meager. The diffraction experiments supply one value of refractive index for each of a

series of electron speeds or wave-lengths, but as these, through errors of measurement or otherwise, do not form a smooth dispersion curve it was not established that the crystal is, in fact, characterized by a parameter which may properly be called its index of refraction for electrons. Our "refractive index" would lose much of the significance which has been attached to it if, for example, it were found to depend not only on wave-length but also upon angle of incidence.

The curve of figure 1 cannot be used to discriminate between the proposed correlations. If the first maximum on the left is the third of the series, the wave-length of the x-ray beam is shorter than that of its electron analogue, but if it is only the second then the wave-length is longer. The former

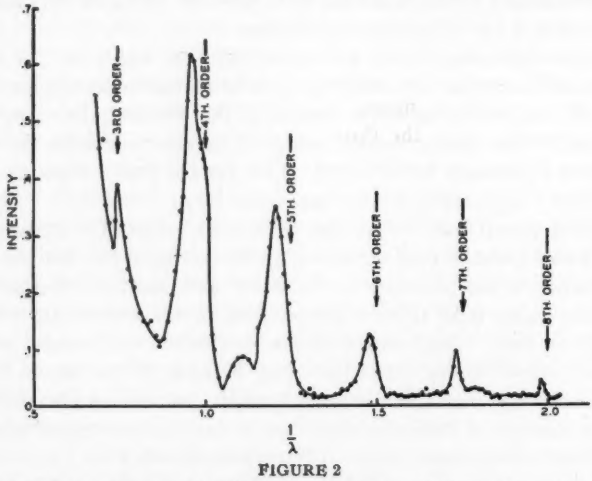


FIGURE 2

Similar to figure 1 under better experimental conditions—Intensity of reflected beam vs. $1/\lambda$ for $\theta = 10°$.

correlation is perhaps suggested more strongly, but the situation is really the same as with the diffraction data. We do not know with certainty whether the analogue of a given electron beam is the adjacent x-ray beam of shorter wave-length or the adjacent beam of longer wave-length. The former correlation leads to values of refractive index greater than unity and the latter to values less than unity. Both sets of values are given in our previous note.

The further observations which have been made lend considerable support, we think, to the view that electron refraction in the optical sense is a property of the crystal, and that the indices are greater than unity. On the other hand the results are less simple than might be expected and more data will be required before a really definite answer can be given. For

bombarding potentials above 150 volts the results so far obtained are consistent with a dispersion curve represented by $\mu = (1 + \varphi/V)^{1/2}$, the value of φ being about 18 volts. But for bombarding potentials below 150 volts and above 65 volts (the limit of our observations) the situation is much more complicated. More than one value of index is found for certain wave-lengths within this region, and it is not yet certain that this peculiarity is independent of the order of the reflection. It appears also that under certain conditions reflection occurs as if the index of refraction were unity; intensity maxima occur, the data of which satisfy the ordinary Bragg formula.

One maximum of this kind is shown very clearly in figure 2, in which we have plotted a second set of observations on the reflection beam for angle of incidence 10 degrees taken under better experimental conditions than those of figure 1. (In figure 2 the intensity has been plotted against the reciprocal of wave-length instead of against the square-root of the bombarding potential, as in figure 1.) The secondary maximum which coincides in position with the third order Bragg beam is indicated also in figure 1, although at the time these earlier observations were made we did not believe in the reality of this feature. There is an indication in figure 2 of a fourth order Bragg beam, and of other minor features, the most prominent of these being the small maximum at $1/\lambda = 1.11$ A^{-1}.

We would like to suppose that the beams which satisfy the Bragg law are due to constructive interference among the waves proceeding from the numerous crystal facets which make up the surface of the target. These facets differ in level by integral multiples of the Bragg constant d, and together constitute what amounts to an echelon grating. That the waves proceed from small lattices associated with the facets rather than from the surface layers of atoms only would seem not to invalidate this explanation.

Observations similar to those of figure 2, but less complete, have been made at other angles of incidence up to $\theta = 52$ degrees. The results of all such measurements are exhibited in figure 3. The dots in this diagram coördinate the electron wave-lengths and the cosines of the angles of incidence at which intensity maxima of the reflected beam have been observed, the area of each dot being roughly proportional to the strength of the maximum. The data corresponding to the maxima of figures 1 and 2 occur at the extreme right (abscissa 0.985). The straight lines are the graph of the Bragg formula $\lambda = (2d/n) \cos\theta$ in its various orders. If the intensity maxima of the electron reflection beam were determined in the same way as are the intensity maxima of x-ray reflection the dots would, of course, fall along these straight lines.

The observation is that the dots fall on lines and curves which display the same general characteristics as the Bragg lines; they have the same trend and are similarly spaced. The suggestion from the figure is that the

dots lying just above the fourth order Bragg x-ray line determine the fourth order electron line, and that the dots which fall between the second and third order Bragg lines correspond to third order reflections. On the other hand, it is conceivable that the orders in these cases are actually the third and second rather than the fourth and third. To discriminate between these possibilities and at the same time to test the hypothesis that electron waves experience ordinary optical refraction, we have calculated and plotted four sets of quantities: The indices of refraction corresponding

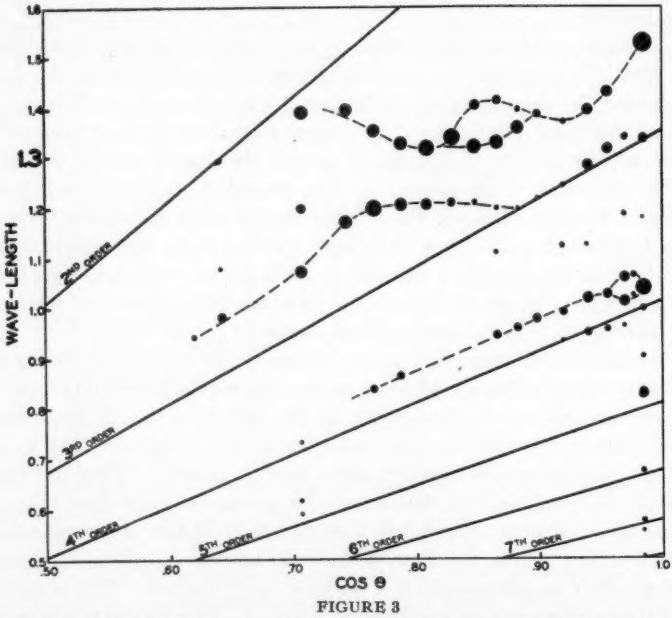


FIGURE 3

The positions of the dots on this figure correlate the wave-lengths and the angles of incidence at which electron reflection occurs from a {111}-face of a nickel crystal. The area of each dot represents the intensity of the reflection.

to the dots in figure 3 assuming the first assignment of orders mentioned above, and then the second, and also the "spacing factors" defined in our *Physical Review* article for the same two assignments. It was anticipated that one at most of these sets of quantities when plotted against wave-length would form a continuous curve. In three of the cases the curves seem to be definitely discontinuous; the curve determined by observations in one order does not join up with the curve of the next order. The remaining case is that of the refractive indices corresponding to the first assignment, and here continuity if not thoroughly established is at least

strongly indicated. The formulas used in calculating indices and "spacing factors" are the following:

$$\mu^2 = n^2\lambda^2/4d^2 + \sin^2\theta, \quad \beta = n\lambda/2d \cos\theta$$

The indices corresponding to the first assignment are plotted in figure 4, a different kind of symbol being used for each order of reflection. It will be noted that at $1/\lambda = 0.93\ A^{-1}$ observations in adjacent orders yield nearly the same value of index, and that this is true also at $1/\lambda = 1.02\ A^{-1}$, at $1/\lambda = 1.06\ A^{-1}$ and at $1/\lambda = 1.20\ A^{-1}$. The solid curve to the right of $1/\lambda = 1.00\ A^{-1}$ is a graph of $\mu = (1 + \varphi/V)^{1/2} = (1 + \varphi\lambda^2/150)^{1/2}$ with the value of φ adjusted to obtain as good a fit as possible with the observed points in this range. The value of φ is 18 volts and the agreement is as good, we think, as can be expected. The graph of this relation is continued to the left by the interrupted curve.

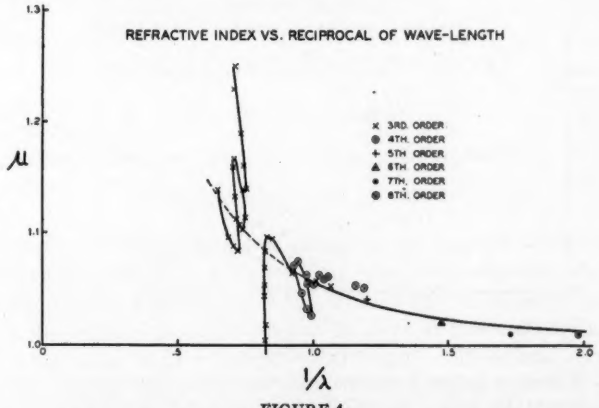


FIGURE 4

μ vs. $1/\lambda$—Refractive indices for nickel calculated from the data of figure 3.

What interpretation is to be placed upon the peculiar behavior of the index in the region below $1/\lambda = 1.00\ A^{-1}$ we do not know. The form of the curve in the region about $1/\lambda = 0.8\ A^{-1}$ suggests some sort of resonance phenomenon such as that to which we ascribe the anomalous dispersion of light. But oscillators and resonance would appear to have no place in the theory of electron refraction. Refraction in the two cases arises from quite different causes, as is evident from the fact that the trend of what is presumably the normal dispersion curve of nickel for electrons is opposite to that of the normal dispersion curves of materials for light and x-rays; in the former case $d\mu/d\lambda$ is positive while in the latter it is always negative.

We have not succeeded in identifying the position of the region of "anomalous dispersion" definitely with any of the known critical constants of nickel. The critical potentials of soft x-ray emission from nickel determined by various experimenters[6] show nothing striking near 90 volts, nor do Petry's values of the critical potentials of secondary electron emission from nickel. The wave-length of the K x-ray absorption limit for nickel, 1.49 A, is fairly near the long wave-length edge of the "anomalous dispersion" region at 1.4 A.

In the data which we have at hand there is an indication that the bifurcation of the curves shown in figure 3, at $\lambda = 1.32\ A$, $\cos\theta = 0.082$, in the third order and at $\lambda = 1.02\ A$, $\cos\theta = 0.96$ in the fourth order, occurs again at $\lambda = 0.82\ A$, $\cos\theta = 0.98$ in the fifth order.

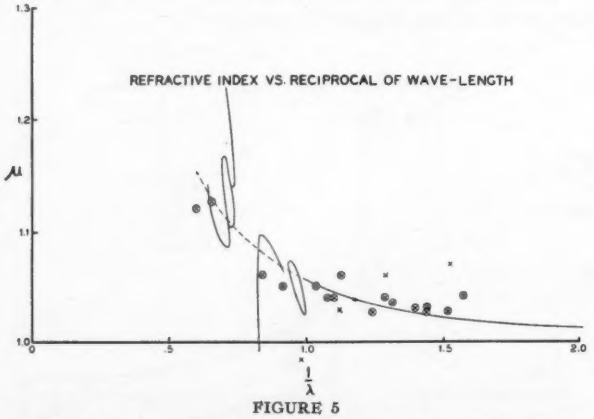


FIGURE 5

Refractive indices calculated from the data of the diffraction experiment (loc. cit.). More reliable points indicated by crossed circles, less reliable points indicated by crosses.

It is possible that the behavior of the reflection beam is not independent of the azimuth of the plane of incidence. The results given here are for the incident beam lying in one of the {111}-azimuths of the crystal as described in our previous note. We hope to look for a dependence upon azimuth at some future time.

In figure 5 we reproduce the dispersion curve of figure 4 and in the same diagram plot values of the refractive index calculated from our earlier observations on the electron diffraction beams. These calculations are, of course, based on the second correlation of electron and Laue beams. The values of μ differ somewhat from those given by Bethe because, in reducing the data, we have in each case applied a small correction to the observed angle of the beam to make wave-length and angle satisfy pre-

cisely the plane grating formula. It will be noted that, with the exception of two, all of the points fall to the right of the regions of "anomalous dispersion," and that none of them falls in this region. It is due to this circumstance presumably that the displacements of the electron diffraction beams from their x-ray analogues display no marked abnormalities. It will be noted also that although the values of μ calculated from the diffraction beams are rather scattered they are not inconsistent with the dispersion curve constructed from the more precise data of the reflection beams.

[1] Davisson and Germer, *Proc. Nat. Acad. Sci.*, **14**, 317 (1928).
[2] Davisson and Germer, *Nature*, **119**, 558 (1927); *Phys. Rev.*, **30**, 705 (1927).
[3] Eckart, *Proc. Nat. Acad. Sci.*, **13**, 460 (1927).
[4] Bethe, *Naturwiss.*, **15**, 787 (1927).
[5] Bethe, *Ibid.*, **16**, 333 (1928).
[6] Andrewes, Davies and Horton, *Proc. Roy. Soc.*, **117**, 660 (1928).

OSCILLATIONS IN IONIZED GASES

BY IRVING LANGMUIR

RESEARCH LABORATORY, GENERAL ELECTRIC CO., SCHENECTADY, N. Y.

Communicated June 21, 1928

In strongly ionized gases at low pressures, for example in the mercury arc, the free electrons have a Maxwellian velocity distribution corresponding to temperatures that may range from 5000° to 60,000°, although the mean free path of the electrons may be so great that ordinary collisions cannot bring about such a velocity distribution. Electrons accelerated from a hot cathode (primary electrons), which originally form a beam of cathode rays with uniform translational motion, rapidly acquire a random or temperature motion which must result from impulses delivered to the electrons in random directions.

In this laboratory we have been studying these phenomena[1] in detail during the last 4–5 years, but the mechanism underlying the Maxwellian distribution and its extremely short time of relaxation have not been understood. At an early date it occurred to me that electric oscillations of very high frequency and of short wave-length in the space within the tube might produce a scattering of the kind observed, but calculation showed that average field strengths of several hundred volts per centimeter would be necessary and this seemed an unreasonable assumption. Experiments capable of detecting oscillations of the electrodes with amplitudes greater than 0.2 volt failed to show such oscillations.

Ditmer[2] although unable to detect oscillations, concluded that oscilla-

tions of frequencies higher than 10^8 probably caused the scattering but was not able to suggest why there should be such oscillations.

Penning[3] detected oscillations of frequencies from 3×10^8 to 6×10^8 per second and found that the electron scattering and the oscillations nearly always occurred together. No cause was assigned for the oscillations.

Dr. Tonks and I have repeated and have confirmed Penning's observations. The amplitude of the oscillations is small, less than 0.2 volt, and frequencies up to 1.2×10^9 have been observed. These waves, which are of well-defined frequency, appear to depend on the presence of the ultimate electrons (low velocity electrons) and can be observed in any part of the bulb. Other oscillations of much lower frequency (2×10^7 to 20×10^7 per second) can be observed in regions transversed by the primary electrons, which are injected into the tube with velocities corresponding to 25 to 70 volts. These oscillations and the scattering of the primary electrons are observable only when the current of primaries is raised to 10 milliamps. or more.

It seemed that these oscillations must be regarded as compressional electric waves somewhat analogous to sound waves. Except near the electrodes, where there are *sheaths* containing very few electrons, the ionized gas contains ions and electrons in about equal numbers so that the resultant space charge is very small. We shall use the name *plasma* to describe this region containing balanced charges of ions and electrons.

For purposes of calculation we may consider the plasma to consist of a continuum of positive electricity having a charge density ρ with free electrons distributed within it, the average electron space charge being $-\rho$. If the electron temperature is zero, the electrons will probably arrange themselves in a face-centered cubic space lattice (close packed arrangement).

Any electron after displacement from its equilibrium position oscillates about that point with a frequency ν which is given approximately by

$$\nu^2 = \frac{e}{m} \cdot \frac{\rho}{3\pi} = \frac{e^2 n}{3\pi m} \qquad (1)$$

where n is the number of electrons per cubic cm., e is the charge and m the mass of the electron. At low temperatures or high electron densities (actually if $T < 0.001\, n^{1/3}$) the plasma is thus a kind of crystalline gas.[4] However the electron temperature will ordinarily be far too high to permit this kind of structure even if it were consistent with the actual discontinuous distribution of positive charge. We must therefore consider that the electrons move rapidly throughout the plasma.

By means of the Boltzmann and the Poisson equations, Debye and Hückel[5] have shown that under equilibrium conditions in a plasma, the average potential near a charged plane varies according to the law

$$V = V_0 \epsilon^{-x/\lambda_D}$$

where x is the distance from the plane, $\epsilon = 2.718$ and λ_D which we may call the Debye distance, is given by

$$\lambda_D = \sqrt{\frac{kT}{4\pi ne^2}} = 6.92 \sqrt{\frac{T}{n}} \text{ cm.} \qquad (2)$$

where T is the temperature of the electrons.

Thus at any distance from the plane which is large compared to λ_D, the average potential and the average electric field become zero. However, when the potential of the plane is changed it requires a certain time for the redistribution of charges to occur and these transient fields are not given by the Debye-Hückel theory.

Plasma Oscillations.—If throughout a volume of dimensions large compared to λ_D we change the concentration of electrons by some transient external means, the resulting electric fields act in such a direction as to equalize the concentration, but the potential energy of these fields is converted into kinetic energy of the electrons so that oscillations occur, and electric waves may result.

Mr. Harold Mott-Smith has worked out the theory for infinite trains of harmonic compressional waves of length λ traveling through a uniform plasma containing electrons of negligible velocity. He finds that the velocity of propagation of the waves v is given by

$$v = \lambda \sqrt{\frac{e}{m} \frac{\rho}{\pi}} \qquad (3)$$

Thus no matter what the wave-length, the frequency (v/λ) in sec.$^{-1}$ is

$$\nu = \sqrt{\frac{e}{m} \cdot \frac{\rho}{\pi}} = 8980 \sqrt{n} \qquad (4)$$

which is $\sqrt{3}$ times that given by equation (1) for the oscillations of single electrons.

Since the velocity is proportional to the wave-length, the waves show high dispersion and the group velocity of the waves turns out to be zero. That is, although these waves can propagate through space with the velocity v they transmit no energy.

Dr. L. Tonks has proved that if the electron density in any region within the plasma is slightly altered so that the potential distributon is $f(x, y, z)$, then after the time t it is $f(x, y, z) \cos 2\pi\nu t$, where ν has the value given by equation (4). This is a stationary wave which remains of constant wave shape and does not spread.

In these derivations it was assumed that the wave-length λ (or the grain size) of the disturbance $f(x, y, z)$ is so large that the plasma electrons do not

move distances comparable to λ during one period $(1/\nu)$ of the oscillations. This condition is satisfied if λ is large compared to the Debye distance λ_D.

Beam Oscillations.—A beam of electrons moving with a uniform translational velocity v and containing n_1 electrons per cm.^{3} may contain similar oscillations of a frequency ν_1 which may be calculated from equation (4) by placing $n = n_1$. In order that these beam oscillations may be unaffected by the plasma electrons it is necessary that their wave-length λ_1 shall be small compared to v/ν, where ν is the frequency of the plasma oscillations. With electron beams of 50 volts energy v/ν is usually about $50\lambda_D$ and thus beam oscillations can exist having a wide range of wave-lengths whose frequency can be calculated by equation (4). Since n_1 is far smaller than n, the beam frequency ν_1 will be less than ν, but because of the Doppler effect the waves will appear to have different frequencies according to the direction from which they are observed. Transversely, the frequency may be ν_1 but longitudinally frequencies as high as v/λ may occur.

The oscillations that we have observed seem to be in general accord with this theory. Measurements of n, the concentration of ions, by means of the positive ion currents and calculation of the frequency ν from this by equation (4) give values that agree reasonably well with the definite frequencies of about 10^9 that are found by using a Lecher system. The broad band of lower frequencies from 10^7 to 10^8 sec.$^{-1}$ which are observed only in the regions traversed by the primary electrons appear to correspond to beam oscillations which are spread over a range in frequencies by the Doppler effect and the non-uniform concentration of the primary electrons.

The smallness of the observed amplitude of the oscillations in the experiments is readily explained by the fact that the plasma oscillations do not transmit energy, or rather, that energy can be transmitted only because of some second order effects neglected in the elementary theory. Thus it is probable that the plasma oscillations may well be of sufficient amplitude to cause the electron scattering that brings about a Maxwellian distribution.

Ionic Plasma Oscillations.—In a recent paper J. S. Webb and L. A. Pardue[6] have described experiments which showed oscillations in low pressure discharges in air with frequencies that ranged from 1 to 240 kilocycles. The frequency increased with the current and with the voltage and decreased in general as the pressure increased. It seems possible that such oscillations might be ionic oscillations in agreement with equation (4) where m is now the mass of the ion. If the electron temperature is very high compared to the ion temperature, as should be expected at high pressure, the diffusion of the electrons will set up fields which tend to make the concentration of ions uniform so that equation (4) should apply.

A more complete analysis by Dr. Tonks, assuming a Maxwellian dis-

tribution of electrons but negligible ion temperature, shows that when the wave-length λ of the ionic oscillations becomes small compared to $2\pi\lambda_D$ (calculated by equation (2) from the electron temperature) the frequency approaches the limiting maximum value given by equation (4), and the group velocity approaches zero. On the other hand, as λ becomes large compared to $2\pi\lambda_D$ the wave velocity and the group velocity both approach the limiting maximum value

$$v = \sqrt{kT_e/m_p} = 3.9 \times 10^5 \sqrt{T_e m_e/m_p} \text{ cm. per sec.}$$

The frequencies v/λ thus decrease as the wave-length increases just as in the case of sound waves. These waves are in fact electric sound waves and their frequencies lie in the range of Webb and Pardue's experiments.

Thermal Equilibrium of Plasma Oscillations.—Any arbitrarily selected volume (of dimensions large compared to λ_D) which we may choose within a plasma may be regarded as a harmonic oscillator, whose damping coefficient will be of the order λ_D/λ. Since electrons, with a velocity distribution corresponding to T_e are continually entering and leaving this selected volume or cell, the oscillator should acquire on the average $^1/_2 kT_e$ of kinetic energy and $^1/_2 kT_e$ of potential energy (electrostatic field) for each degree of freedom. The kinetic energy of the oscillator is included in that of the electrons which coöperate to produce the oscillation but the potential energy is a new factor, not included in the ordinary theory of electron motions.

If we have N independent cells per unit volume the total energy density of the electric field of the oscillations is $^3/_2\, NkT_e$; equating this to $X^2/8\pi$, where X is the field intensity, gives

$$\left.\begin{aligned} X^2 &= 12\pi NkT_e \\ \text{or } X &= 2.16 \times 10^{-5}\sqrt{NT_e} \text{ volts per cm.} \end{aligned}\right\} \quad (5)$$

If there were no lower limit for the wave-length of harmonic plasma oscillations, we could put $N = n$ the concentration of electrons, as in the Debye theory of the specific heats of solids. We have seen that the potential distribution $f(x, y, z)$ in any plasma oscillator is arbitrary. Thus to find N we might expand $f(x, y, z)$ into a Fouriers' series, breaking off the series when we have reached the permissible grain size: the number of independent parameters thus fixes N. Perhaps a better way of attacking the problem is to assume that there are n independent oscillators per unit volume but those of shorter wave-length are non-harmonic and are highly damped so that for these the time average of the potential energy may be far less than the average kinetic energy.

For the present we can obtain an approximate value of N by assuming that harmonic oscillators are possible down to the wave-length $\lambda = \alpha\lambda_D$,

where α is a small numerical factor that probably lies between 2 and 10.

Thus $N = 1/(\alpha\lambda_D)^3$ and from equation (5) and (2) the field strength X is

$$X^2 = 96\pi^{5/2}\alpha^{-3}(kT)^{-1/2}n^{3/2}e^3$$

or

$$X = 1.17 \times 10^{-6}\alpha^{-3/2}n^{3/4}T^{-1/4} \text{ volts per cm.} \quad (6)$$

For a typical low pressure gas discharge in which $n = 10^{11}$ and $T = 10^4$ this gives $X = 4$ volts per cm. if we assume $\alpha = 3$. This is the mean field intensity of plasma oscillations in thermal equilibrium with the ultimate electrons. According to equation (4) the frequency would be 2.8×10^9 per second, corresponding to that of a radio wave of wave-length 10 cm.

Electric Fields Due to a Random Distribution of Electrons.—Let us compare these fields produced by the plasma oscillations with those to be expected from a random distribution of electrons. We assume provisionally that the positive charge in the plasma is continuous, and that the probability per unit volume for the occurrence of electrons is constant throughout the plasma. The radius of a sphere which contains on the average one electron is

$$r_1 = (3/4\pi n)^{1/3} = 0.62\ n^{-1/3} \quad (7)$$

The average field X_1 at the center of the sphere averaged for all positions of the electron is

$$X_1 = 3e/r_1^2 = 1.12 \times 10^{-6}n^{2/3} \text{ volts cm.}^{-1} \quad (8)$$

The effect of all the other electrons (out to $r = \infty$) increases the average field only about 10.3%. Thus the average field X_a at any point in space is

$$X_a = 1.23 \times 10^{-6}n^{2/3} \text{ volts cm.}^{-1} \quad (9)$$

This field is always stronger than that of the plasma oscillations, as given by equation (6). In our example in which $n = 10^{11}$, we find $X_a = 26.5$ volts per cm. as compared with our estimate of 4 volts per cm. for the plasma oscillations.

Most of the effects produced by the local fields depend upon the square of the field strength so that we are usually interested in the root-mean square field which we shall denote by X_m. If the electrons are regarded as point charges the value of X_m at any fixed point is infinite. In the neighborhood of any given electron, however, according to the Boltzmann equation the average concentration of electrons is

$$n\epsilon^{-a/r}$$

where

$$a = e^2/kT = 1.66 \times 10^{-3}/T \text{ cm.} \quad (10)$$

We thus find that the mean field X_m which acts on any electron because of the presence of all the others is given by

$$\left.\begin{aligned} X_m^2 &= 4\pi nkT \\ \text{or} \qquad X_m &= 1.25 \times 10^{-5}\sqrt{nT} \text{ volts cm.}^{-1} \end{aligned}\right\} \quad (11)$$

from which $X_m = 396$ volts per cm. if $n = 10^{11}$, $T = 10^4$.

The mean field X_r due to all electrons that lie at a distance greater than r (when $r \gg a$) is found to be

$$\left.\begin{aligned} X_r^2 &= 4\pi ne^2/r \\ \text{or} \qquad X_r &= 5.08 \times 10^{-7}\sqrt{n/r} \text{ volts cm.}^{-1} \end{aligned}\right\} \quad (12)$$

If we put $r = \lambda_D$ as given by equation (2)

$$X_D = (4\pi ne^2)^{3/4}(kT)^{-1/4} = 1.93 \times 10^{-7} n^{3/4} T^{-1/4} \text{ volts cm.}^{-1} \quad (13)$$

Comparing this with equation (6) which gives the field due to plasma oscillations we see that the two equations are identical if $\alpha = 3.31$ which is a reasonable value. We may conclude that the fields due to plasma oscillations in thermal equilibrium with the electrons are of about the same magnitude as the fields acting at any point originating from all electrons at distances greater than λ_D. We have assumed, however, that the electrons are distributed at random, that is, as if they were uncharged. According to Debye-Hückel, the charges cause the distribution to be more uniform so that the field due to electrons at distances greater than λ_D should disappear. The field of the plasma oscillations thus take the place of the field that is forbidden by the Debye-Hückel theory. The definite frequency characteristic of these oscillations makes them, however, essentially different from the irregularly fluctuating fields that they replace.

Energy Delivered to Electrons by the Field.—If a steady electric field of intensity X acts for a time τ on an electron which is initially at rest, the electron acquires the energy

$$E = X^2e^2\tau^2/2m$$

The fields acting on a given electron resulting from a random distribution of neighboring electrons fluctuate rapidly in direction and magnitude. If τ is the time of relaxation of these fields we may consider roughly that during each interval τ the field is steady, and of magnitude X, but in each interval the direction is independent of that in the preceding or following intervals. In a time t which is large compared to τ the total energy acquired by an electron is then

$$E = X^2e^2\tau t/2m \quad (14)$$

The field produced by electrons at a distance r will fluctuate at a rate such that, approximately

$$\tau = r/v$$

where v, the velocity of the electrons, is

$$v = \sqrt{8kT/\pi m} = 6.23 \times 10^5 \sqrt{T} \text{ cm. sec.}^{-1}$$

By integrating (14) over all values of r from 0 to r we find that the rate W_r at which energy is acquired by a given electron is

$$W_r = \pi^{3/2} e^4 n (2mkT)^{-1/2} \log_e \left(\frac{rkT}{1.78e^2}\right) \text{ erg. per second.} \qquad (15)$$

By dividing this into $\frac{3}{2} kT$ we find the time of relaxation of the electrons, that is, the time required for an electron to gain or lose the average energy. Multiplying this time by the velocity v we find an approximate value for λ_e, the free path of an electron. This free path, which may be defined as the average distance an electron travels before losing most of its energy of motion in that direction, is thus

$$\lambda_e = \frac{6}{\pi^2} \frac{(kT)^2}{e^4 n} \frac{1}{\log_e \left(\frac{rkT}{1.78e^2}\right)}$$

or

$$\lambda_e = \frac{9.5 \times 10^4 T^2}{n \log_{10} (r/1.78a)} \qquad (16)$$

where a is given by equation (10). We have seen from equation (12) that the field produced by the more distant electrons is small compared to that due to the nearer electrons. However, because of the fact that the field of the distant electrons fluctuates more slowly, this field becomes important in determining the energy transfer between electrons. Thus, according to equation (15) the rate of energy transfer increases indefinitely as r approaches infinity. It is clear that equations (12) and (15) are applicable only for values of r that are not much greater than λ_D since the field of electrons at a greater distance fluctuates periodically at a frequency which no longer decreases as r increases (plasma oscillations).

In our example $n = 10^{11}$, $T = 10^4$ and thus by equation (2) $\lambda_D = 2.2 \times 10^{-3}$ cm. Putting these in equations (10) and (16) we find $a = 1.66 \times 10^{-7}$ and the free λ_e is 24.6 cm. The experimental data obtained with mercury arcs shows effective free paths very much shorter than this. With a 0.1 ampere arc in a 3-cm. tube containing mercury vapor saturated at 15° C. we found[7] $T = 36000°$ and $n = 3.4 \times 10^9$ and yet the Maxwellian distribution, which was destroyed at the tube walls by the removal of the fastest electrons, re-established itself within one centimeter, so that λ_e must have been less than 1 cm. But under these conditions equation (2) gives $\lambda_D = 2.3 \times 10^{-2}$ cm. and by equations (10) and (16) $a = 4.6 \times 10^{-8}$ cm. and $\lambda_e = 6600$ cm. Thus the fields due to a random distribution of electrons are not capable of accounting for the observed short free path.

The fields due to the plasma oscillations also deliver energy to the electrons. The oscillations of large wave-length, being harmonic and only slightly damped, deliver very little energy, but the plasma oscillations of the shortest wave-length, $\alpha\lambda_D$, are highly damped and the electrons which are being acted on are moving distances comparable to $\alpha\lambda_D$ during a single period of the oscillations. Thus as a rough approximation we may consider the time of relaxation τ of the field X to be equal to $1/2\nu$, a half period. The rate at which energy is delivered is then by equations (14) and (4)

$$W_0 = X^2e^2\tau/2m = 1/4X^2e\sqrt{\pi/mn} \qquad (17)$$

Assuming that the oscillations are in thermal equilibrum with the electrons, X is given by equation (6) so that equation (17) becomes

$$W_0 = \frac{24\pi^3\alpha^{-3}e^4n}{(mkT)^{1/2}} \qquad (18)$$

Comparing this with the rate at which energy is delivered by the field of electrons distributed at random, equation (15), we find

$$W_0/W_r = \frac{82.1}{\alpha^3 \log_{10}(r/1.78\,a)} \qquad (19)$$

As before we may identify r with λ_D; then if $n = 10^{11}$ and $T = 10^4$ we find that $W_0 = W_r$ if $\alpha = 2.8$. Thus plasma oscillations in thermal equilibrium, although their fields are much weaker than those produced by the nearer electrons, contribute about as much as the latter to the energy interchanges between electrons. Any complete theory of the kinetics of a plasma cannot ignore the plasma oscillations.

Plasma Oscillations of Greater Amplitude.—To account for the experimentally observed electron-scattering the amplitude of the plasma oscillations must be far greater than that which corresponds to thermal equilibrium with the electrons. Then, too, with thermal equilibrium, the maximum energy of the electrical oscillations in a coupled Lecher system would be kT whereas actually the oscillations that Penning and we have observed are far more intense than this.

In a mercury arc or a discharge with a hot cathode, ions are continually being generated in a random manner throughout the plasma, and are diffusing to the walls of the tube and to the anode where they recombine with electrons. The electrons set free by the ionization process (secondary electrons) usually have far higher energies than the ultimate electrons in the plasma so that the secondaries are rapidly separated from the ions which were simultaneously produced. Thus the ionization tends to create a random distribution, whereas the electric field in accord with the Debye-Hückel theory, tends to make the field more uniform. In this way energy

is delivered to the plasma oscillations, particularly to those of large wave-length, so that the amplitudes become larger than those corresponding to thermal equilibrium.

In a plasma containing n electrons per unit volume, let us consider a sphere of volume V and radius r. With a uniform distribution of electrons there would be nV, which we may call q_0, electrons in the sphere. But with a random distribution the actual number q will differ from q_0 by an amount whose average value is $\sqrt{q}$. Thus the average charge of the sphere is $\pm e\sqrt{q}$. Assuming this charge to be uniformly distributed within the sphere, the total electrostatic energy of the field (of which $^5/_6$ lies outside the sphere) is

$$E = (3/5)qe^2/r = (4\pi/5)e^2r^2n \tag{20}$$

and the field at the surface of the sphere is

$$\left.\begin{aligned} X_r^2 &= (4\pi/3)ne^2/r \\ \text{or} \\ X_r &= 2.9 \times 10^{-7}\sqrt{n/r} \text{ volts per cm.} \end{aligned}\right\} \tag{21}$$

If the sphere is of large size compared to λ_D the electrostatic energy given by equation (20) becomes the potential energy of a spherical plasma oscillation which is only slowly converted into thermal energy. Let us compare the energy E of equation (20) with the energy $^1/_2kT$ which is the average potential energy of the plasma oscillation in thermal equilibrium. Expressing kT by equation (2) in terms of λ_D we find

$$\frac{E}{^1/_2kT} = \frac{2}{5}\left(\frac{r}{\lambda_D}\right)^2$$

We see that the energy associated with a sphere of radius r is equal to $^1/_2kT$, the value required for thermal equilibrium, if $r = 1.58\lambda_D$, but this energy increases in proportion to r^2. Thus, if $n = 10^{11}$, $T = 10^4$, $\lambda_D = 2.2 \times 10^{-3}$ and we had initially a random distribution of electrons, the energy of the plasma oscillations associated with a sphere of 3 cm. radius would be 2×10^6 times greater than if there were thermal equilibrium.

This may possibly account for the amplitude of the observed oscillations detected by the aid of a Lecher system, but it hardly seems able to explain the observed scattering which must depend on the plasma oscillations of shorter wave-length, for these are not so far from the conditions of thermal equilibrium.

We are now investigating other possible causes that may produce plasma oscillations of large amplitude.

[1] Langmuir, *Phys. Review*, **26**, 585 (1925); *Zeit. Physik*, **46**, 271 (1927).

[2] Ditmer, *Phys. Review*, **28**, 507 (1926).

[3] Penning, *Nature*, Aug. 28 (1926), and *Physica*, **6**, 241 (1926).

[4] The possibility of gaseous crystals in stellar atmospheres has been recently considered by Eddington, *Monthly Notices, Roy. Astron. Soc.*, **88**, 369 (1928).

[5] Debye and Hückel, *Phys. Zeitsch.*, **24**, 185, 305 (1923).

[6] *Bull. Amer. Phys. Soc.*, **3**, 19 (1928).

[7] Langmuir and Mott-Smith, Jr., *General Electric Review*, **27**, 766 (1924). Table XIV.

EVIDENCE THAT THE COSMIC RAYS ORIGINATE IN INTERSTELLAR SPACE

BY ROBERT A. MILLIKAN AND G. HARVEY CAMERON

NORMAN BRIDGE LABORATORY OF PHYSICS, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated July 12, 1928

If it may be regarded as established by the evidence heretofore advanced[1] that the cosmic rays are the signals sent out through the heavens of the creation of the common elements out of positive and negative electrons, the next important question to attempt to answer is "where are these creative processes going on?" To this question there are two different sorts of possible answers, as follows:

(1) In the stars where pressures, densities and temperatures may, one or all, be enormously high, or else

(2) In interstellar space where pressures, densities and temperatures are all extraordinarily low.

In both of these localities matter exists under extreme and as yet unexplored conditions, and in view of the history of the last thirty years of physics, it would no longer be surprising if matter were again found to behave in some hitherto unknown and unexpected way as a new field of observation is entered.

Of the two foregoing alternatives we think it possible to eliminate the first and to establish the second with considerable definiteness, and that for the two following reasons.

First.—If the mere presence of matter in large quantities and at high temperatures favored in any way the atom-building processes which give rise to the cosmic rays, then it is obviously to be expected that the sun, in view of its closeness, would send to the earth enormously more of them than could any other star. But the fact is that all observers are agreed that the change from midday to midnight does not influence at all the intensity of the cosmic rays. This can only mean that *the conditions existing in and about the sun, and presumably also in and about other stars as well, are unfavorable to the atom-building processes which give rise to these rays.*

Since, however, the rays do come to us at all times, day and night, and, according to all observers, at least very nearly equally from all directions—according to some, as accurately as they have as yet been able to make the

measurements—there is scarcely any escape from the conclusion that *the atom-building processes giving rise to the cosmic rays are favored by the conditions existing in interstellar space.* If then, in going from a point in interstellar space toward the center of a star the favorable conditions for atom-building existing in outer space have disappeared as the surface of the star is reached, it is well-nigh inconceivable that they will again reappear in penetrating from the surface to the center—a path along which the changes in physical conditions all continue unchanged in direction. So that from the foregoing we may not only conclude quite definitely that the stars are not the sources of the cosmic rays, but also that the main *atom-building processes probably do not take place inside of stars at all.*

Second.—The foregoing conclusions may also be arrived at from an entirely different mode of approach, namely, from our measurements upon the absorption coefficients and the total energy content of the cosmic rays.

The hardest rays which we have observed are completely absorbed (reduced to say 2% of their initial intensity) in going through 70 meters of water. This means that, even if the atom-building processes went on inside a star, the resulting cosmic radiations could not possibly get out, but would all be frittered away in heat* before emergence, save in the case of those rays that originated in the star's very outermost skin—a skin equivalent in absorbing power to a hundred or so meters of water.

But we have also found that the total energy coming into the earth's atmosphere in the form of cosmic rays is about one-tenth the total heat and light energy coming to the earth from the stars exclusive of the sun.[3] This last fact means that *if the cosmic rays have their origins within the stars they cannot, even at the points of their origin, have an intensity more than ten times that which they have when they reach the earth's atmosphere, for if they had then the cosmic-ray energy transformed into heat by absorption on the way out would yield a total heat outflow from the stars larger than the observed ten to one ratio.* In other words, if the stars are the sources of the observed cosmic rays, it follows from our measurements on absorption coefficients and on total energy content that the total heat output of the stars must be furnished by the atom-building processes going on in their merest outer skins of a thickness equivalent in absorbing power to about a hundred meters of water, and that therefore no atom-building processes, nor any other activities capable of furnishing heat, can then be going on in their interiors.

It is, however, so altogether absurd to suppose that atom-building processes are going on actively at the surface of a star, and down to a depth of a hundred meters, and then suddenly stop there, that we are forced back by this present mode of approach to the same conclusion arrived at from the direct determination of the lack of cosmic-ray activity of a particular star, the sun, namely, to the conclusion that *the observed cosmic rays do not origi-*

nate in the stars at all, but that they must originate under the extreme influences of exactly the opposite sort existing in interstellar or intergalactic space.

These considerations bring us then from two entirely new points of view to the conclusion that the heat output of the stars must be derived from an entirely different source from the atom-building processes which produce the cosmic rays. Jeans[4] and Eddington,[5] from other considerations based wholly upon the lifetimes of the stars, have repeatedly emphasized the necessity of finding a source for this output other and greater than the process of atom-building, but we can now go further and say that the process of energy emission by atom-building does not take place in the stars at all, or at least in such amount as to make the stars an appreciable factor in the output of cosmic rays, for if it did the star would have to be radiating heat much faster than is the case. As is well known, Eddington and Jeans have found this new source of stellar heat *not in an atom-building process, but rather in an atom-annihilating process* which they assume to be going on in the interior of stars, positive electrons being thought to be continually transforming their entire mass into ether waves in accordance with the demands of Einstein's equation. As indicated above, we have sought in vain among our cosmic rays for a ray of penetrating power corresponding to this act. It will be recalled that the mass which disappears in the creation out of hydrogen of one gram-atom of silicon—this produced the hardest cosmic ray that we can say with certainty that we have yet observed, for the iron rays are still to some degree hypothetical—was 0.23 g. The complete annihilation of the mass of hydrogen would obviously then produce, in accordance with Einstein's equation, a ray having approximately 4 times (accurately $\frac{1.0778}{0.23}$ times) the energy and penetrating power of our hardest definitely observed ray. *Our failure to find this ray, however, is no argument at all against the existence of the process in the interior of stars where the pressures are colossal and the densities may be enormous.* Indeed, our failure to find this ray means rather that, if the act occurs at all, as Eddington and Jeans think it must, it is obliged to occur precisely in the interior of stars where the resulting radiation is hidden away behind an impenetrable screen of matter—a screen that transforms all its energy into heat before the ray can get out. If the cosmic rays originated within the stars they would of course be similarly screened.

On the other hand, that the atom-*building* processes responsible for the cosmic rays, as distinct from the atom-*destroying* process just considered, actually occur, as our experiments definitely show, *outside the stars*, or at least where the rays produced by them can get to us, and in an energy that is of the same order of magnitude as that of the heat poured out by the star, is an extraordinarily illuminating fact. For it suggests at once, when combined with Eddington's argument, the following incom-

plete cycle, each element in which now has the *experimental* credentials indicated in the brackets:

(1) Positive and negative electrons exist in great abundance in interstellar space (see the evidence of the spectroscope).

(2) These electrons condense into atoms under the influence of the conditions existing in outer space, viz., absence of temperature and high dispersion (see the evidence of the cosmic rays).

(3) These atoms then aggregate under their gravitational forces into stars (see the evidence of the telescope).

(4) In the interior of stars, under the influence of the enormous pressures, densities and temperatures existing there, an occasional positive electron, presumably in the nucleus of a heavy atom, falls into complete coincidence with a negative, i.e., transforms its entire mass into an ether pulse the energy of which, when frittered away in heat, maintains the temperature of the star and furnishes most of the supply of light and heat which it pours out (see the evidence of the lifetimes of the stars—Eddington-Jeans).

The foregoing is as far as the experimental evidence enables us to go; but the recent discovery of the second element of the above unfinished cycle, namely, that the supply of positive and negative electrons is being used up continually in the creation of atoms, the signals of whose birth constitutes the cosmic rays, at once raises imperiously the question as to why the process is still going on at all after the æons during which it has apparently been in process—or better *why the building stones of the atoms have not all been used up long ago.* And the only possible answer seems to be to complete the cycle, and to assume that these building stones are continually being replenished throughout the heavens by the condensation, with the aid of some as yet wholly unknown mechanism, of radiant heat into positive and negative electrons.

This is a new mode of approach to a conclusion a portion of which at least is old. For the Einstein assumption itself, that mass is convertible into radiant energy, requires the existence also of the inverse process, unless the validity of the Second Law of Thermodynamics, in the form of the principle of microscopic reversibility, is to be denied. The effort to work out the thermodynamics of a cycle containing the Einstein process, but without sacrificing microscopic reversibility, has recently been made by Stern,[6] Tolman,[7] and Zwicky.[8]

But we have in the foregoing gone farther than they. *The essentially new element that we have introduced is the experimental observation that the creative, or atom-building, processes do not appear to take place at all in the stars, but only in interstellar or intergalactic space, where densities and temperatures are practically zero.* Our *experimental* evidence does not, indeed, extend to the creation of the lightest element, hydrogen, out of radiation,

but the inclusion of this also among the creative processes going on only in interstellar space, is a natural extension of our observational data on the other abundant elements. For making this extension we are denying the reversibility at high temperatures and pressures of the process of the transformation of matter into radiation. This is why our conclusion differs from that of Stern, Tolman and Zwicky, and why we are able to regard the universe as in a steady state now, although a state not satisfying the condition of microscopic reversibility.

* It is important to remember that, as we have already shown, rays[2] of this kind become frittered away into heat in this passage through matter *without any change in the quality* (i.e., frequency or absorption coefficient) *of the residual beam.*

[1] Millikan, R. A., and Cameron, G. H., *Proc. Nat. Acad. Sci.*, **14**, 445 (1928).

[2] Millikan, R. A., and Cameron, G. H., *Phys. Rev.*, **28**, 851 (1926).

[3] Millikan, R. A., and Cameron, G. H., *Phys. Rev.*, **31**, 921 (1928).

[4] J. H. Jeans, "Problems of Cosmogony and Stellar Dynamics," Cambridge, **1919**, p. 286.

[5] A. S. Eddington, "The Internal Constitution of the Stars," Cambridge, **1926**, Chap. XI.

[6] Stern, O., *Zeits. Electrochemie*, **31**, 448 (1925).

[7] Tolman, Richard C., *Proc. Nat. Acad. Sci.*, **12**, 67 (1926); **14**, 268, 348, 353 (1928).

[8] Zwicky, F., *Proc. Nat. Acad. Sci.*, **14**, August (1928).

THE DISPLACEMENT INTERFEROMETRY OF BAROMETRIC PRESSURE

BY CARL BARUS

DEPARTMENT OF PHYSICS, BROWN UNIVERSITY

Communicated June 30, 1928

1. *Apparatus.*—The use of the interferometer U-gauge as a microbarometer reading well within 10^{-4} cm. would seem to be easy of accomplishment; for the apparatus is very simple and the readings sharp and smooth. Actually the investigation is made extremely difficult, since the pressure gauge is also a very sensitive air thermometer. I shall in the following paragraphs show this purposely under varying atmospheric temperature.

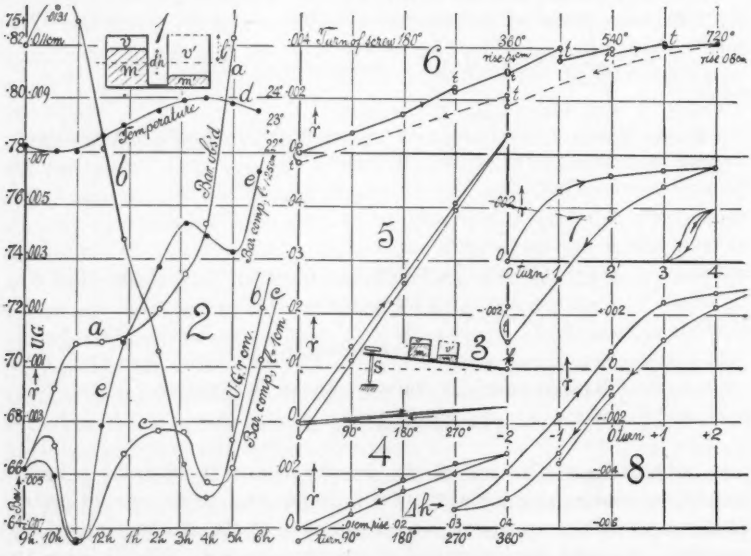
A diagram of the apparatus is given in figure 1, in which the shank of the U-gauge, vm, is closed and the other $v'm'$ open to the atmosphere. The mercury mm' is contained in a massive block of iron and the cisterns are about 10 cm. in diameter. As the heads to be observed are relatively small, they may be treated as differentials, $dh = dr \cos\theta$, being the mercury head obtained from the micrometer displacement dr (for an angle of incidence $\theta = 45°$ of the rays to the mirrors), necessary to bring the achro-

matic fringes of the quadratic interferometer back to zero. The equation holds for finite values of h and r.

This dh is the first pressure at v in question. A second one results from the compression of the volume v of air. This may be taken as $-h(dv/v) = h(dh/2)$, if h is the height of the barometer and $l = 1$ cm. the initial depth of the volume v. The third pressure results from a rise of temperature dt of the air in v and may be computed as hdt/t, t being the absolute temperature. Thus the full increment Δh, of the barometer, if the level in v is above that in v', is

$$\Delta h = dh + h(dh/2l + dt/t)$$

When the level v is below v', dh is negative as in the following experiments. Usually h may be taken as the mean height of the barometer.



It is seen at once that the quantity in parenthesis is critical, owing to the large factor h, so that it is not uncommon to obtain a persistently negative dh, even in case of a rising barometer (Fig. 2).

2. *Data.*—Figure 2 is an example of results for a rising barometer (curve a) read to 0.01 cm. and reduced to 20° C. The change of temperature in the room is shown in curve d, read off a thermometer placed near the U-gauge, but outside of it. Curve b shows the corresponding micrometer reading r at the interferometer by which the fringes are brought back to the initial zero, r_o, so that the effective micrometer displacement will be $r - r_o$. This curve is quite definite and an error of 10^{-4} cm. in r which

would only just be visible in the figure is improbable, even with the small fringes (about 0.01 cm. in the ocular) used. A difference of height of the mean level of the mercury surfaces even within 1 mm. should be recognized (3 fringes), if atmospheric pressure only were in question.

It is therefore a little disconcerting to find the calculated values of the barometer resulting in the wildly meandering curve c. This curve, computed for a gauge depth $l = 1$ cm. is too low as a whole. Another computation was therefore made for $l = 1.25$ cm. and the result is shown in curve e. The effective l is hard to estimate and could probably best be obtained from equation (1) itself, if dt/t is guaranteed.

The oscillations of curves c and e are obviously temperature discrepancies since $^1/_{10}$ degree centigrade is equivalent to 0.015 cm. of the barometer. In the morning the massive block of iron is likely to be colder at the outset than the atmosphere, and the temperature of the former rises over a larger range; whereas in the afternoon the reverse is the case. This rise moreover is probably steady, so that all atmospheric fluctuations result in spurious corrections. I did not in the present work care to dismantle the interferometer, to sink a thermometer into the body of the gauge; but this is clearly necessary, even if the conditions of practically constant temperature are guaranteed as in the summer. Curve e in figure 2 is not so wild as some others obtained, because curve d is here more free from the usual temperature indentations, each of which can be identified (sometimes with a lag and enormously amplified) in the computed barometer graphs.

Finally it is not probable that the zero of the interferometer can be guaranteed for 10 hours or more, so that the variation here must be tested.

3. *Time Interval Reduced.*—If the time of observation is short so that temperature changes vanish, the behavior of the gauge is more satisfactory. To test this the gauge was placed in the middle of an incline of variable angle (see insert Fig. 3) about 50 cm. long. This had a knife edge at k and a graduated screw S (0.182 cm. per turn) at the other end, by which the inclination was changed positively or negatively from zero. Since the distance between the centers of the mercury pools was 11 cm. these were raised $\Delta h = (11/50) \times 0.182 = 0.0400$ cm. relative to each other, for each turn of the screw.

The gauge was first tested with both shanks open to the atmosphere. In this case there should be no displacement of fringes with the inclination, since the glass plates on top of the gauge pools compensate each other. To my surprise a small but definite fringe displacement, quite regular (as shown in figure 4), did indeed occur. The abscissas show the mean elevation of one mercury reservoir (solid) above the other. The rise of the mercury level is somewhat less than $^1/_{10}$ of this, the ordinates being Δr where $\Delta h = 0.71\ \Delta r$. The Δh gradually decreases and the return curve shows hysteresis.

It is difficult to account for this except as due to surface rigidity and surface viscosity of the mercury, remembering that the pools undergo considerable oscillation at each adjustment. The mercury film sustains a pressure here exceeding $\Delta h = 0.0020$ cm. of mercury and the backlash is about 0.0004 cm. of mercury.

One shank of the gauge was now closed to the atmosphere and the same experiment performed with the results shown in figure 5. The mercury thus rises nearly as much as the solid case of the gauge. There is also hysteresis and its value is about as large as the full deflection in figure 4 (see lower curve in the same scale). Nevertheless, the graphs in figure 5 are appreciably straight.

We may write $dp/p = dh/h = dv/v = dl/l$, where $h = 76$ cm. and $l = 1.7$ cm. From figure 5, dr per turn is 0.05375 whence $dh = dr \times 0.707 = 0.0380$ cm. Hence $dl = 0.00085$ cm. which registers half the linear expansion of the air in the closed shank. Since dh in the inclined plane test was 0.0400 cm. per turn and in the fringe test 0.0380 cm., the difference, 0.0020 cm., agrees as well with $2\,dl = 0.0017$ cm. as can be expected from these rather trying experiments with liquid surfaces.

The rise of the gauge as a whole, per turn of screw, would be $(1/2) \times 0.182 = 0.091$ cm., which is equivalent to a fall of the barometer of $0.091 \times 9 \times 10^{-4} = 8.1 \times 10^{-5}$ cm. This alone would be equivalent to nearly 3 fringes; but as it is associated with the rise of the mercury surface of $dh = 0.0380$ cm.. from reduced air pressure in the closed shank, it is but 0.002 of the observed fringe displacement, and thus necessarily lost in experiments of the present kind.

4. *Hysteresis Loops of Surface Viscosity.*—The phenomenon of figure 4 with the shanks of the U-gauge both open to the atmosphere deserves further study. Figure 6 shows that the rise of mercury level ($\Delta h = \Delta r \cos 45°$) or pressure supported by capillary forces, practically ceases when the screw has made two turns. Each turn lifts the center of one surface $\Delta h = 0.04$ cm. above the other. Furthermore, though vigorous tapping changes the level somewhat (marked t in the diagram), it does not wipe out the phenomenon. On the return passage it actually accentuates it as the arrows indicate.

In figure 7 similar results are given for whole turns of the screw, each producing a differential lift of surface centers of 0.04 cm. It is obvious that here is an interesting beginning of a hysteresis loop of the kind suggested by the insert. The return curve is particularly striking. Vigorous tapping (t) at the end wipes out but a part of the negative residual surface strain. The completion of a cycle on both the positive and negative side is very difficult; for the fringes are sooner or later distorted by an incipient lack of parallelism in the mercury surfaces.

In a later repetition of the cycle with four turns, all endeavors to get the

negative feature failed; but a cycle (not shown) of two turns of screw, positive and negative, was completed. After the first cycle the succeeding repetitions reproduced the outer points as accurately as the fringes could be set. In figure 8 the same experiment is completed with 3 turns of the screw, positive and negative. In figure 8 and others, the lower branches are not quite symmetrical to the upper branches; but this is no doubt due to the fringes as suggested, unless a residual surface strain masks the true zero.

THE STRUCTURE OF ACETYLENE

BY JARED KIRTLAND MORSE

DEPARTMENT OF PHYSICS, UNIVERSITY OF CHICAGO

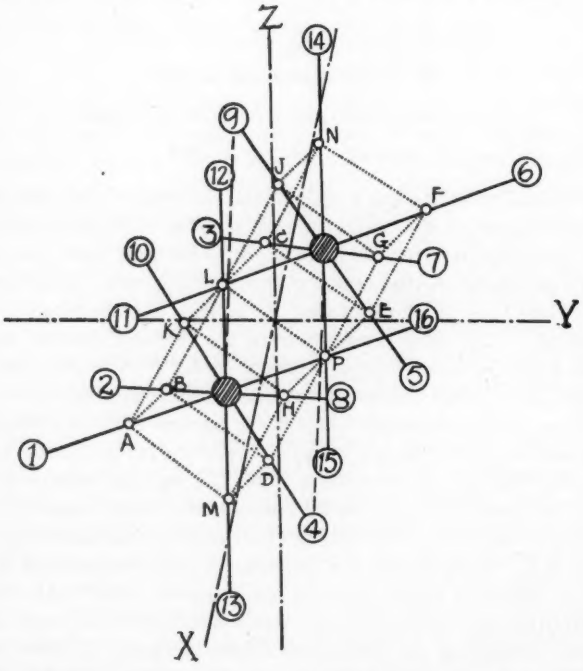
Communicated July 9, 1928

In a previous investigation a scale model for ethane[1] was presented and all the coördinates of the carbon nuclei, possible hydrogen positions and electron positions, referred to three rectangular axes were computed in terms of a parameter R (the radius of the carbon atom). The value of this parameter had been previously found[2] to be very close to three quarters of an angstrom. The value obtained from the diamond lattice being 0.77 A.U. and from Bernal's graphite lattice 0.75 A.U. The ideas underlying this model are (1) that carbon atoms can be conveniently represented by cubes whose center represents the carbon nucleus and the corners the L electron positions, (2) that the union between the carbon atoms in ethane consists of a two electron bond (cube edge), (3) that the union between the carbon atoms and hydrogen nuclei consists of a single electron bond in which the two nuclei and the binding electron lie on the same straight line. This type of C—H bond allows of the location of sixteen possible hydrogen positions of which in ethane only six are occupied. This scale model has been quantitatively verified[3] by building a number of them into a crystal lattice and comparing the computed diffraction patterns with those obtained by x-ray analysis of solid ethane. Furthermore, these same ideas have been applied to methane[4] in which there are eight possible hydrogen positions and since there are only four hydrogen nuclei available a certain finite number of arrangements are possible, conditioned by a number of ways in which four points can be distributed among the eight corners of a cube. The moments of inertia of some of these "dynamic isomers" of methane have been computed on the basis of the model, and these calculated moments of inertia agree quantitatively with those obtained from the analysis of the infra-red band spectra of methane.

Since acetylene can be considered as an ethane molecule from which four

hydrogen atoms have been removed and since the infra-red bands of acetylene have recently been analyzed by Levin and Meyer,[5] it is now possible to write down the coördinates of the dynamic isomers of acetylene, compute the moments of inertia of these structures and compare them with the moments of inertia obtained from the band spectra analysis.

Figure 1 represents a drawing of the acetylene molecule. The carbon nuclei are represented by large cross-latched circles, the possible L electron positions by small white circles lettered A, B, C, D, E, F, G, H, J, K, L, M, N and P, and the cubical nature of the carbon atoms is indicated



by dotted lines. The possible hydrogen positions are indicated by large white circles numbered from (1) to (16) inclusive. The C—H bonds are indicated by solid lines and the coördinate axes X, Y and Z are represented by dashes and dots.

Table 1 gives the coördinates of all the elements of this structure in terms of the radius of the carbon atom (half the body diagonal of the cube formed by the electron positions) referred to three rectangular axes passing through the center of the molecule.

Since in the model herewith presented there are sixteen possible hydrogen

positions all linked to their respective carbon atoms by precisely the same type of bond and since there are only two hydrogen nuclei available to fill these sixteen positions it is obvious that a certain finite number of unique structures are possible. A consideration of a three-dimensional model of

TABLE 1

COÖRDINATES OF THE ACETYLENE MOLECULE IN TERMS OF R (RADIUS OF THE CARBON ATOM)

	x	y	z
		Carbon Nuclei	
	0.000	+0.472	+0.667
	0.000	−0.472	−0.667
		Possible Hydrogen Positions	
(1)	0.000	−2.357	−1.333
(2)	+1.633	−1.414	0.000
(3)	+1.633	−0.472	+1.333
(4)	+1.633	+0.472	−1.333
(5)	+1.633	+1.414	0.000
(6)	0.000	+2.357	+1.333
(7)	−1.633	+1.414	0.000
(8)	−1.633	+0.472	−1.333
(9)	−1.633	−0.472	+1.333
(10)	−1.633	−1.414	0.000
(11)	0.000	−1.414	0.000
(12)	0.000	−0.472	+1.333
(13)	0.000	−0.472	−2.667
(14)	0.000	+0.472	+2.667
(15)	0.000	+0.472	−1.333
(16)	0.000	+1.414	+1.333
		Possible Electron Positions	
(A)	0.000	−1.414	−1.000
(B)	+0.817	−0.944	−0.333
(C)	+0.817	0.000	+1.000
(D)	+0.817	0.000	−1.000
(E)	+0.817	+0.944	+1.000
(F)	0.000	+1.414	+1.000
(G)	−0.817	+0.944	+0.333
(H)	−0.817	0.000	−1.000
(J)	−0.817	0.000	+1.000
(K)	−0.817	−0.944	−0.333
(L)	0.000	−0.472	+0.333
(M)	0.000	−0.472	−1.667
(N)	0.000	+0.472	+1.667
(P)	0.000	+0.472	−0.333

this structure indicates that a number of these structures have a similar shape and lie in a plane. One of these having hydrogen positions occupying the positions numbered (9) and (4) is indicated in figure 1 by broken lines. From the coördinates of Table I the moments of inertia of any

structure can be easily calculated. The following list represents the simplest and most symmetrical of some of the possible dynamic isomers of acetylene. The numbers indicate the positions of the two hydrogen nuclei.

1. *Dynamic Isomers* (*1*) (*6*) *and* (*13*) (*14*).—Both carbon and hydrogen nuclei lie in a plane. The largest moment of inertia of this figure is about an axis through the center of mass perpendicular to this plane and if the mass of the hydrogen atom is taken to be 1.66×10^{-24} gm. the moment of inertia is easily computed from the coördinates given in table 1 using a value of $R = 0.77 \times 10^{-8}$ cm. to be 3×10^{-39} gm. cm.^{2}

2. *Dynamic Isomers* (*10*) (*5*), (*2*) (*7*), (*4*) (*9*), *and* (*8*) (*3*).—These also are plane figures and have a similar configuration but slightly smaller dimensions. The value of their maximum moment of inertia is 2.5×10^{-39} gm. cm.^{2}

3. *Dynamic Isomers* (*11*) (*16*), *and* (*12*) (*15*).—These also are plane figures having a similar configuration and their maximum moment of inertia is 2.0×10^{-39} gm. cm.^{2}

Levin and Meyer[5] analyzed bands in the infra red in the regions 13.7μ 7.5μ, and 3.0μ. The average $\Delta\nu$ for these respective regions were 2.35 cm.$^{-1}$, 2.35 cm.$^{-1}$, and 2.34 cm.$^{-1}$, which leads on the basis of the quantum theory to the moment of inertia $I = 2.4 \times 10^{-39}$ gm. cm.^{2} by means of the relation

$$I = \frac{h}{4\pi^2 c(\Delta\nu)} \qquad (1)$$

and the same constant on the basis of the classical theory has the values for these respective regions of 2.3, 2.1 and 2.0×10^{-39} gm. cm.^{2}

It will be noted that the moment of inertia of 2.5×10^{-39} gm. cm.^{2} which represents the moment of inertia of the most abundant of the dynamic isomers calculated on the basis of the model herewith presented, from absolute values of the radius of the carbon atom (diamond) and the length of the (C—H) bond (solid ethane) obtained from x-ray experiments is in remarkable agreement with the value 2.4×10^{-39} gm. cm.^{2} obtained by Levin and Meyer[5] by means of the quantum theory from the analysis of the infra-red absorption bands of gaseous acetylene. Such an agreement for a single compound might well be a coincidence, but when it is remembered that a similar agreement between the observed and calculated moments of inertia was obtained from methane,[4] and that the lattice built of molecular models of ethane[1] gave diffraction effects in quantitative agreement with those actually obtained by x-ray analysis[2] and that in all cases these models were based on the same simple postulates and the same values of the fundamental constants were employed throughout, such success ceases to be a coincidence and assumes a fundamental importance in the field of molecular structure.

The simplicity of these fundamental postulates and the ease in which they can be applied and verified by experiment open up alluring possibilities for their application throughout the whole range of organic compounds and seem to present a most valuable clew for extending our knowledge of complex molecular structures.

Besides the intrinsic interest of the results herewith presented on acetylene this work offers a convenient point of departure for (1) the interpretation of the x-ray diffraction effects from solid acetylene and the determination of the lattice structure, (2) the study of the relative abundance of the various dynamic isomers through an analysis of the effect of temperature on the intensities of these bands, and (3) a study of the possible electron configurations in the acetylene molecule and the physical effects to be looked for from them. Since there are fourteen electron positions, and since there are only ten electrons available to fill them, the unsaturated character of acetylene is accounted for on the model, and different allocations of the available electrons in these possible electron positions allow for the occurrence of different electronic structures which may be called "electromers" of acetylene.

[1] J. K. Morse, "The Structure and Dimensions of the Ethane Molecule," *Proc. Nat. Acad. Sci.*, **14**, 37–40 (1928).

[2] J. K. Morse, "Atomic Lattices and Atomic Dimensions," *Proc. Nat. Acad. Sci.*, **13**, 227–32 (1927).

[3] J. K. Morse, "The Lattice Structure of Ethane," *Proc. Nat. Acad. Sci.*, **14**, 40–5 (1928).

[4] J. K. Morse, "The Structures of Methane," *Proc. Nat. Acad. Sci.*, **14**, 166–71 (1928).

[5] A. Levin and C. F. Meyer, "The Infra Red Absorption Spectra of Acetylene, Ethylene and Ethane," *J. Optical Soc. Am.*, **16**, 137–64 (1928).

ON DERIVATIVES OF NON-ANALYTIC FUNCTIONS

BY E. R. HEDRICK

UNIVERSITY OF CALIFORNIA AT LOS ANGELES

Communicated July 2, 1928

1. *Introduction.*—It is my purpose in this paper to discuss the relations that exist between the recent papers by Kasner[1] and the paper by Ingold, Westfall, and myself,[2] regarding the derivatives of functions of a complex variable which are not analytic. These we have called *non-analytic;* Kasner has called them *polygenic*, but he also uses the name non-analytic. Some of the ideas of these papers evidently agree, but there are some which do not agree on the face of things. At one point, indeed, I shall introduce a notation different from that of our former paper, in order to make clear

the intimate connection that exists, but I shall point out that these changes do not affect the essential character of the concepts discussed by us previously.

Kasner has (loc. cit.) considered the values of the limit of the increment-ratio

$$\gamma = \alpha + i\beta = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z} \tag{1}$$

of a non-analytic (or polygenic) function $w = u + iv = f(z)$ of the complex variable $z = x + yi$. Using his notation, $u = \phi(x, y)$, $v = \psi(x, y)$, we have

$$\gamma = \alpha + i\beta = \frac{\phi_x + i\psi_x + m(\phi_y + i\psi_y)}{1 + im} \tag{2}$$

where m denotes the slope dy/dx of the path along which Δz approaches zero, and where γ depends in general on m. He shows that the locus of γ for various values of m is a circle:

$$\left(\alpha - \frac{\phi_x + \psi_y}{2}\right)^2 + \left(\beta - \frac{-\phi_y + \psi_x}{2}\right)^2 = \left(\frac{\phi_x - \psi_y}{2}\right)^2 + \left(\frac{\phi_y + \psi_x}{2}\right)^2, \tag{3}$$

whose radius is zero if and only if the Cauchy-Riemann equations

$$\left.\begin{aligned} \phi_x &= \psi_y, \\ \phi^y &= -\psi_x \end{aligned}\right\} \tag{4}$$

hold at the point (x, y). This circle Kasner calls the *derivative circle*. I shall call it the *Kasner circle*.

2. *Diameters of the Kasner Circle.*—Kasner has used the fact that the point

$$\gamma = \alpha + i\beta = \psi_y + i\psi_x$$

lies on the circle (3), and has given a simple construction for determining the point on the circle that corresponds to any slope m. It is obvious that the points that correspond to $m = 0$ and to $m = \infty$, that is,

$$\frac{dw}{dx} = \phi_x + i\psi_x, \text{ and } \frac{dw}{dy} = \psi_y - i\phi_y \tag{5}$$

also lie on the Kasner circle, and that they are at the ends of a diameter of it. These points give an easy construction, and they furnish a good mnemonic for the center and the radius of the circle.

Another diameter of the circle is furnished by finding the maximum and the minimum distances from the origin in the (α, β) plane to the circle. This we did in our paper (loc. cit.). We set

$$r = \alpha^2 + \beta^2 = \left|\frac{dw}{dz}\right|^2 = \frac{du^2 + dv^2}{dx^2 + dy^2}, \tag{6}$$

from which, after easy cancellations, we found

$$r = \left|\frac{dw}{dz}\right|^2 = \frac{E + m^2(E + G') + m^4G + 2m(1 + m^2)F}{(1 + m^2)^2} = \frac{E + 2mF + Gm^2}{1 + m^2}, \quad (7)$$

where E, F and G are defined by the usual equations

$$E = \phi_x^2 + \psi_x^2, \qquad G = \phi_y^2 + \psi_y^2, \qquad F = \phi_x\phi_y + \psi_x\psi_y.$$

We found that the maximum and the minimum of r are given by the equation

$$F + (G - E)m - Fm^2 = 0, \quad (8)$$

and that they are the solutions of the quadratic*

$$\rho^2 - (E + G)\rho + J^2 = 0. \quad (9)$$

3. *The Indicatrix.*—In our paper, we used the ratio r, rather than its square root, for convenience; and we defined two concepts, each of which would not be affected materially by such a change. Calling the solutions of the equation (9) ρ_1 and ρ_2, we called the ellipse whose semi-axes are ρ_1 and ρ_2, the *indicatrix*, and we called the difference

$$|\rho_1 - \rho_2| \quad (10)$$

the *ellipticity*, as a measure of the departure of the ellipse from circularity. Since we were concerned chiefly with the directions of the axes, which are the solutions of (8), we chose the ellipse and the measure of ellipticity in as simple a manner as possible. Any other simple measure based on ρ_1 and ρ_2 would have served our purpose as well; for example, the difference

$$d = \sqrt{\rho_1} - \sqrt{\rho_2} \quad (11)$$

is an equally satisfactory measure. *This difference d is precisely the diameter of the Kasner circle.* I propose henceforth to call d the *ellipticity*.

A similar change of notation is desirable in the case of the indicatrix, in order to bring it into close relation with the ideas discussed above. As we defined it, the directions were the solutions of (8), and the semi-axes were ρ_1 and ρ_2. There exists another ellipse with the same principal directions, whose radial distances bear closer relations to the preceding concepts. For, let us plot as polar distances from the original values (x, y), the *stretching factor* P defined by the equations

$$P^2 = \frac{1}{r} = \frac{ds^2}{d\sigma^2} = \left|\frac{dz}{dw}\right|^2 = \frac{dx^2 + dy^2}{du^2 + dv^2}, \quad (12)$$

where s and σ are the arc-lengths in the (x, y) and the (u, v) planes, respec-

tively. Let us set $\bar{x} = x - x_0$, $\bar{y} = y - y_0$, and lay off P in the direction $m = dy/dx$. Then we shall have

$$P^2 = \bar{x}^2 + \bar{y}^2, \; m = \bar{y}/\bar{x},$$

and (12) becomes

$$E\bar{x}^2 + 2F\bar{x}\,\bar{y} + G\bar{y}^2 = 1, \tag{13}$$

which is an ellipse, since

$$J^2 = \begin{vmatrix} \phi_x\phi_y \\ \psi_x\psi_y \end{vmatrix}^2 = H^2 = EG - F^2 > 0, \tag{14}$$

provided, of course, that J is not zero at (x_0, y_0).

Likewise, if J is not zero at (x_0, y_0), and with mild assumptions of continuity, we can solve for x and y in terms of u and v near (x_0, y_0), and we may write

$$x = \Phi(u, v), \quad y = \Psi(u,v), \tag{15}$$

$$\mathfrak{E} = \Phi_u^2 + \Psi_u^2, \; \mathfrak{F} = \Phi_v^2 + \Psi_v^2, \; \mathfrak{G} = \Phi_u\Phi_v + \Psi_u\Psi_v, \tag{16}$$

$$\mathfrak{F}^2 = \mathfrak{E}\mathfrak{F} - \mathfrak{G}^2 = \frac{1}{EG - F^2} = \frac{1}{J^2}. \tag{17}$$

With this notation, we may draw, in the (u, v) plane, about the point (u_0, v_0), an ellipse whose polar distances are the values of R, where

$$R^2 = r = \frac{d\sigma^2}{ds^2} = \left|\frac{dw}{dz}\right|^2 = \frac{du^2 + dv^2}{dx^2 + dy^2} = \frac{du^2 + dv^2}{\mathfrak{E}du^2 + 2\mathfrak{F}dudv + \mathfrak{G}dv^2}. \tag{18}$$

The equation of this ellipse, in terms of $\bar{u} = u - u_0$, and $\bar{v} = v - v_0$, is

$$\mathfrak{E}\bar{u}^2 + 2\mathfrak{F}\bar{u}\,\bar{v} + \mathfrak{G}\bar{v}^2 = 1. \tag{19}$$

The ellipse (19) has as its radial distances precisely the distances from the origin in the (α, β) plane to the corresponding points of the Kasner circle, and the difference between its semi-axes is the diameter of the Kasner circle. This ellipse is, in fact, the polar diagram of the Kasner circle. The difference of its semi-axes, in the older notation, is

$$\sqrt{\rho_1} - \sqrt{\rho_2},$$

which is the diameter of the Kasner circle. It is evident that the changes made above are not material from the standpoint of the earlier paper, since the concepts discussed there, particularly the concepts called *principal directions, characteristic lines, analytic at a point, functions of the same class*, etc., would be the same under either set of notations. For the purpose of displaying the relationships mentioned above, however, it is advantageous to use the stretching ratios R and P defined above, which are expressed in terms of the original ratio r by the simple formulas

$$R^2 = r, \quad P^2 = \frac{1}{r};$$

and it is advantageous to substitute the ellipses (13) and (18) in the (x, y) and the (u, v) planes, respectively, for the somewhat arbitrary ellipse used in the original paper.

4. *The Jacobian of the Increment-Ratio.*—The increment-ratio itself,

$$\zeta = \xi + i\,\eta = \frac{\Delta w}{\Delta z} = \frac{w - w_0}{z - z_0} = \frac{f(z) - f(z_0)}{z - z_0} = F(z), \quad (20)$$

defines a function $F(z)$ of the complex variable z, except for $z = z_0$. For an analytic function, or for a non-analytic one that has an isolated analytic point at $z = z_0$, the derivative exists uniquely, and its value completes in a continuous manner the definition of $F(z)$ at $z = z_0$. If this derivative vanishes, however, there are in general two leaves of a Riemann surface which connect together about the point $F(z_0)$, and the rate of angular turn about that point is in general doubled.

When $f(z)$ is not analytic at $z = z_0$, some (but not all) of the properties that depend upon the derivative in the analytic case can be stated in terms of the jacobian

$$J = \frac{D(\xi, \eta)}{D(x, y)} = \begin{vmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{vmatrix}.$$

It is not difficult to calculate this jacobian, but I shall omit it here for brevity. It turns out, very simply, that the locus of the points at which the jacobian vanishes, that is, the locus of the equation

$$J = \frac{D(\xi, \eta)}{D(x, y)} = 0,$$

is precisely the Kasner circle in the (ξ, η) plane, which coincides with the (α, β) plane of Kasner's paper insofar as these points are concerned. I may therefore state the theorem that follows.

THEOREM. *The equation of the Kasner circle is* $J = 0$.

In general, the vanishing of the jacobian for any non-analytic function (that is, for any point transformation of the plane) defines a curve of which the branch-point of an analytic function is a degenerate case, since the jacobian for an analytic function vanishes if and only if the derivative vanishes. For analytic functions, this can happen only at isolated points, since, by the Cauchy-Riemann equations, the jacobian reduces to a sum of squares for an analytic function. For a non-analytic function, however, the locus of the points at which the jacobian vanishes is in general a curve.

The corresponding curve in the (u, v) plane may be called the *edge of regression*, on account of its similarity to the edge of regression of a developable surface. Indeed, the Riemann surface for such a function resembles decidedly a developable surface, since curves in the z plane through any point of it correspond to curves that are tangent to the corresponding curve in the (u, v) plane.

The edge of regression of the increment-ratio $\zeta = \Delta w/\Delta z$ is then precisely the Kasner circle. It is the envelope of curves corresponding to any pencil of curves through the point $z = z_0$ in the (x, y) plane. The Riemann surface over the ζ plane is in general two-leaved, joining along the points of the Kasner circle. The fact discovered by Kasner that the rate of rotation on the Kasner circle is double the rate in the (x, y) plane is a generalization of the fact that the rate of rotation for an analytic function about a branch-point is double the rate about the corresponding critical point.

I shall discuss these concepts in more detail elsewhere.

* There is a misprint in this equation in our original paper. It occurs as equation (10) on p. 335; in that equation the constant term should be J^2.

[1] See E. Kasner, *Science*, **66**, 581–582 (1927); and these PROCEEDINGS, **14**, 75–82 (1928).

[2] See Hedrick, Ingold and Westfall, "Theory of Non-Analytic Functions of a Complex Variable," *J. Math.*, Ser. (9), **2**, 327–342 (1923).

RADIATION AND RELATIVITY. II[1]

BY G. Y. RAINICH

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN

Communicated June 28, 1928

1. In the first paper[2] we arrived at an electromagnetic field which—we are led to believe—is likely to accompany a light corpuscle, and we found that it possesses some properties we are accustomed to associate with light, but it is not periodical. From an entirely different point of view we can arrive at a field in which periodicity appears, so to say, automatically, and although it is not clear at present whether the two points of view can be combined successfully it is probably not without interest to take up this second point of view.

2. It has been proved several years ago[3] that in general relativity theory when the curvature tensor is given the electromagnetic tensor is determined (up to a constant) and that the expressions for the electromagnetic tensor involve trigonometric functions. It was concluded from this that the electromagnetic field *might* be periodic even when the curvature field is not. But an example of such a situation was lacking; we cannot give an ex-

ample of a non-periodic curvature field with a periodic electromagnetic field now either but we can give an illustration of an analogous situation in flat space. To explain this we may say that in general relativity the curvature field gives the stress-energy tensor which in turn determines (up to a constant) the electromagnetic force-field; it is in the transition from the stress-energy tensor to the force-tensor that the trigonometric functions come in and it is here that we might expect the appearance of periodicity; but this transition from the stress-energy tensor to the force-tensor we may consider also in flat space-time, i.e., in special relativity, or even in classical electrodynamics; and here the situation, although from the point of view that interests us essentially the same, is much easier to handle. *We can actually exhibit cases in which a non-periodic stress-field determines a periodic force-field.* This might, or might not have a bearing on the theory of light but the possibility itself seems not to be devoid of interest, the more so that the statement on the face of it seems to be, if not exactly paradoxical, at least unexpected.

3. Of course, our contention could be proved by exhibiting one special case; but we prefer to indicate the way by which we succeeded in obtaining such cases because it might be used to obtain diversified examples and, possibly, some of them will be found more useful than some others. The pertinent results stated in the note quoted[3] may be formulated as follows: given a field of stresses (i.e., the Maxwellian stresses, the Poynting vector, and the energy), the field of forces which produces these stresses (or is produced by them) may be written as

$$\begin{aligned} X &= a.\cos\phi - f.\sin\phi & L &= a.\sin\phi + f.\cos\phi \\ Y &= b.\cos\phi - g.\sin\phi & M &= b.\sin\phi + g.\cos\phi. \\ Z &= c.\cos\phi - h.\sin\phi & N &= c.\sin\phi + h.\cos\phi \end{aligned} \tag{1}$$

where a, b, c, f, g, h are determined by the stresses and ϕ must satisfy the following differential equations:

$$\begin{aligned} &\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} + \frac{\partial f}{\partial t} = h\frac{\partial\phi}{\partial y} - g\frac{\partial\phi}{\partial z} - a\frac{\partial\phi}{\partial t} \qquad \frac{\partial h}{\partial y} - \frac{\partial g}{\partial z} - \frac{\partial a}{\partial t} = -c\frac{\partial\phi}{\partial y} + b\frac{\partial\phi}{\partial z} - f\frac{\partial\phi}{\partial t} \\ &\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x} + \frac{\partial g}{\partial t} = f\frac{\partial\phi}{\partial z} - h\frac{\partial\phi}{\partial x} - b\frac{\partial\phi}{\partial t} \qquad \frac{\partial f}{\partial z} - \frac{\partial h}{\partial x} - \frac{\partial b}{\partial t} = -a\frac{\partial\phi}{\partial z} + c\frac{\partial\phi}{\partial x} - g\frac{\partial\phi}{\partial t} \\ &\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y} + \frac{\partial h}{\partial t} = g\frac{\partial\phi}{\partial x} - f\frac{\partial\phi}{\partial y} - c\frac{\partial\phi}{\partial t} \qquad \frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} - \frac{\partial c}{\partial t} = -b\frac{\partial\phi}{\partial x} + a\frac{\partial\phi}{\partial y} - h\frac{\partial\phi}{\partial t} \\ &\frac{\partial f}{\partial x} + \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} = -a\frac{\partial\phi}{\partial x} - b\frac{\partial\phi}{\partial y} - c\frac{\partial\phi}{\partial z} \qquad \frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} = f\frac{\partial\phi}{\partial x} + g\frac{\partial\phi}{\partial y} + h\frac{\partial\phi}{\partial z} \end{aligned} \tag{2}$$

(we are not interested here in the restrictions these equations impose on the

stresses—we assume that the stresses are such that the derivatives of ϕ are uniquely determined by (2)). Our purpose is to find periodic force-fields with non-periodic stresses. The stresses do not involve ϕ; the forces as given by (1) contain ϕ under the sin and cos signs. We would obtain periodicity of the character usually associated with light if we could make ϕ take the form $k(t-x)$ with k a constant (the x-axis being in the direction of propagation). This means that we have to seek solutions of the equations (2) for which

$$\partial\phi/\partial y = \partial\phi/\partial z = 0, \quad \partial\phi/\partial t = -\partial\phi/\partial x = k. \tag{3}$$

To simplify the problem we make the additional assumption that a, b, c, f, g, h are independent of x and t. The equations (2) become

$$\begin{aligned} \frac{\partial c}{\partial y} - \frac{\partial b}{\partial z} &= -ka & \frac{\partial h}{\partial y} - \frac{\partial g}{\partial z} &= -kf \\ \frac{\partial a}{\partial z} &= k(h-b) & \frac{\partial f}{\partial z} &= -k(g+c) \\ -\frac{\partial a}{\partial y} &= -k(g+c) & -\frac{\partial f}{\partial y} &= -k(h-b) \\ \frac{\partial g}{\partial y} + \frac{\partial h}{\partial z} &= ka & \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z} &= -kf, \end{aligned} \tag{4}$$

and it is easy to write down the general solution of these equations:

$$\begin{aligned} a &= 2k(q_z - p_y) & f &= -2k(q_y + p_z) \\ b &= 2k^2q - q_{zz} + p_{yz} & g &= -2k^2p + q_{zy} - p_{yy} \\ c &= 2k^2p + q_{zy} - p_{yy} & h &= 2k^2q + q_{zz} - p_{yz}; \end{aligned} \tag{5}$$

here p and q are two harmonic functions in y and z but we must not take two conjugate functions if we want to avoid the exceptional case mentioned in the note.[3] Using these expressions in the formulas (1) and putting $\phi = k(t-x)$ we obtain a field of the required properties.

4. In particular, using in (5) for p and q quadratic harmonic functions we obtain

$$a = 2k[(B-\alpha)y - (A+\beta)z + D - \gamma]$$

$$b = \beta + A + 2k^2\left[A\frac{y^2-z^2}{2} + Byz + Cy + Dz\right]$$

$$c = B - \alpha + 2k^2\left[\alpha\frac{y^2-z^2}{2} + \beta yz + \gamma y + \delta z\right]$$

$$f = 2k[-(A+\beta)y + (B-\alpha)z + C + \delta]$$

$$g = B - \alpha + 2k^2\left[\alpha\frac{y^2 - z^2}{2} + \beta yz + \gamma y + \delta z\right]$$

$$h = -(\beta + A) + 2k^2\left[A\frac{y^2 - z^2}{2} + Byz + Cy + Dz\right].$$

Although a, b, c, f, g, h are not force components it may be interesting to compare the above expressions with the forces obtained from an entirely different point of view in the preceding paper. There is a certain analogy but it is not clear how the two kinds of expressions might be identified. The comparison suggests, however, that we might have to consider a, b, c, f, g, h as variable with $t - x$ and that the expressions (6) would be asymptotically approached by them as $t - x$ approaches zero.

We may note that there is a certain anisotropy which might suggest an interpretation in connection with polarization; but this polarization belongs to the field of stresses rather than to the field of forces and it is of the elliptic type.

[1] Presented before the Chicago meeting of the American Mathematical Society, Christmas, 1926; *Bull. Amer. Math. Soc.*, **33**, 157 (1927).

[2] These PROCEEDINGS, **14**, June, 1928, pp. 484–88.

[3] *Ibid.*, **10**, 294–8 (July, 1924). Also *Trans. Amer. Math. Soc.*, **27**, 106–36 (1925).

CONCERNING PLANE CLOSED POINT SETS WHICH ARE ACCESSIBLE FROM CERTAIN SUBSETS OF THEIR COMPLEMENTS[1]

BY GORDON T. WHYBURN

DEPARTMENT OF PURE MATHEMATICS, UNIVERSITY OF TEXAS

Communicated November 30, 1927*

A point set K is said to be accessible from a point set R provided that if A and B are any two points of K and R, respectively, then there exists a simple continuous arc AB from A to B such that $AB - B$ is a subset of R. Schoenflies[2] has shown that if K is a closed and bounded point set which separates the plane into just two domains D_1 and D_2 and is accessible from each of these domains, then K is a simple closed curve. In this paper a closely related theorem will be established, i.e., it will be shown that if K is any closed and bounded point set in a plane S such that there exist in $S - K$, two mutually exclusive connected point sets R_1 and R_2 such that K is accessible from each of these sets, then either K is a simple closed curve or there exists a simple continuous arc which contains K. With the aid of this proposition it will be shown that if K is any *irreducible cutting*[3] of a plane bounded continuous curve M between two points A and B of M such

that K is accessible from each of the components[4] R_a and R_b of $M-K$ containing A and B, respectively, then either K is a simple closed curve or there exists a simple continuous arc which contains K but which contains no other point of M. A number of other theorems will be given which either lead up to these results or follow readily from them.

A subset K of a connected point set M is said to be a *cutting*[3] of M or is said to cut M provided the set of points $M-K$ is not connected; a cutting K of M is said to be an *irreducible cutting*[3] *of* M provided no proper subset of K is a cutting of M. A subset K of a connected set M is said to be a *cutting of* M *between the points* A *and* B *of* M,[3] or to cut M between A and B, provided that $M-K$ is the sum of two mutually separated point sets M_a and M_b containing A and B, respectively; K is said to be an *irreducible cutting of* M *between* A *and* B[3] provided that K cuts M between A and B but no proper subset of K cuts M between A and B.

The points, or point sets, A and B of a connected point set M are said to be separated in M by a subset N of M provided $M-N$ is the sum of two mutually separated point sets containing A and B, respectively. Wherever it is simply stated that "A and B are separated by a point set N," it is understood that A and B are separated in the plane by the set N. If R is an open subset of a continuum M, then by the M-boundary of R is meant the set of all those points of $M-R$ which are limit points of R. If M and N are point sets, the notation $M \cdot N$ will be used to denote the set of points common to M and N.

THEOREM 1. *If, in a plane* S, R_1, R_2 *and* R_3 *are mutually exclusive connected point sets, then* $S-(R_1+R_2+R_3)$ *contains not more than two points, each of which is accessible from each of these sets.*

Proof.—Suppose, on the contrary, that there exist three points X, Y, and Z in $S-(R_1+R_2+R_3)$ each of which is accessible from each of the sets R_1, R_2 and R_3. Then it readily follows that there exist points A and B in R_1 and R_2, respectively, and arcs AXB, AYB and AZB from A to B, no two of which have in common any point except A and B and such that the set of points $AX+AY+AZ-(X+Y+Z)$ belongs to R_1 and the set $BX+BY+BZ-(X+Y+Z)$ belongs to R_2. Of the three arcs AXB, AYB and AZB, one must lie except for the points A and B within the simple closed curve formed by the sum of the other two. The various cases are alike, so let us suppose $AZB-(A+B)$ lies within the curve $AXBYA$. Let D_1 denote the exterior of this curve, and let D_2 and D_3 be the interiors of the simple closed curves $AXBZA$ and $AYBZA$, respectively. Since R_3 is connected, it is clear that it must lie wholly in one of the domains D_1, D_2, D_3. But this is impossible, for Z is not a limit point of D_1, Y is not a limit point of D_2, and X is not a limit point of D_3, and by hypothesis each of the points X, Y and Z must be a limit point of R_3. Thus the supposition that theorem 1 is not true leads to a contradiction.

The hypothesis of theorem 1 can be weakened somewhat, since in the proof we use only the fact that all of the points X, Y and Z are accessible from R_1 and R_2 and are limit points of R_3.

THEOREM 2. *If R_1, R_2 and R_3 are mutually exclusive arcwise connected subsets of a continuous curve M every subcontinuum of which is a continuous curve, then not more than two points of $M-(R_1+R_2+R_3)$ can be limit points of each of the sets R_1, R_2 and R_3.*

Proof.—Suppose, on the contrary, that there exist three points of $M-(R_1+R_2+R_3)$ each of which is a limit point of each of the sets R_1, R_2 and R_3. But by a theorem of the author's[5] every limit point of R_i ($i=1, 2, 3$) is accessible from R_i, and by theorem 1, not more than two points of $M-(R_1+R_2+R_3)$ can be accessible from each of the sets R_1, R_2, and R_3. This contradiction proves theorem 2.

COROLLARY. *If R_1, R_2 and R_3 are mutually exclusive connected open subsets of a continuous curve M every subcontinuum of which is a continuous curve having M-boundaries B_1, B_2 and B_3, respectively, then $B_1 \cdot B_2 \cdot B_3$ contains not more than two points.*

THEOREM 3. *If every point of the closed and bounded point set K in a plane S is accessible from each of two mutually exclusive connected subsets R_1 and R_2 of $S-K$, then either K is a simple closed curve or there exists a simple continuous arc which contains K.*

Proof.—Suppose K is not a simple closed curve. Let t denote any component of K which contains more than one point. Then t is not a simple closed curve. For suppose it is. Let I and E, respectively, denote the interior and exterior of t. Then neither I nor E can contain both of the sets R_1 and R_2. For suppose E contains both R_1 and R_2; then every point of t is accessible from each of the three mutually exclusive connected subsets R_1, R_2, and I of $S-t$; but this contradicts theorem 1. A like contradiction is obtained if we suppose I contains both R_1 and R_2. Hence it follows that R_1 lies wholly in one of the sets I and E, and R_2 lies wholly in the other. And since every point of K is a limit point of each of the sets R_1 and R_2, then K must be a subset of t, and hence must be identical with t. But t is, by supposition, a simple closed curve, and by our original hypothesis K is not a simple closed curve. Thus the supposition that t is a simple closed curve leads to a contradiction.

Since t is not a simple closed curve, then[6] it contains two points X and Y such that $t-(X+Y)$ is connected. Let Z be any other point of t. Then since, by hypothesis, every point of K is accessible from each of the sets R_1 and R_2, it readily follows that there exist points A and B of R_1 and R_2, respectively, and simple continuous arcs AXB, AYB and AZB such that the point set $AX+AY+AZ-(X+Y+Z)$ is a subset of R_1 and the set $BX+BY+BZ-(X+Y+Z)$ is a subset of R_2. The set of points $t-(X+Y)$, since it is connected, must lie wholly either within or without

the simple closed curve $AXBYA$. The two cases are practically alike, so let us suppose $t - (X + Y)$ lies in the interior I of the curve $AXBYA$. Then the arc AZB divides I into two domains D_1 and D_2 whose boundaries contain X and Y, respectively, and each of which, it is easily seen, must contain points of $t - (X + Y)$. Then if S_1 denotes the set of points $D_1 \cdot t + X$ and S_2 the set $D_2 \cdot t + Y$, it is easily seen that S_1 and S_2 are mutually separated sets and that $S_1 + S_2 = t - Z$. Hence t is disconnected by the omission of any one of its points except X and Y, and therefore[6] t is a simple continuous arc from X to Y.

I shall now show that Z is not a limit point of $K - t$. Suppose the contrary is true. Now the segment $t - (X + Y)$ lies either within or without the curve $AXBYA$. Suppose it lies in the interior I of $AXBYA$. The arc t divides I into two domains G_1 and G_2 whose boundaries contain A and B, respectively. Then either G_1 or G_2 must contain a point P of K. But if P lies in G_1 it is not accessible from R_2, and if P lies in G_2 it is not accessible from R_1; and by hypothesis every point of K is accessible from each of the sets R_1 and R_2. Thus, in this case, the supposition that any point Z of t except X and Y is a limit point of K leads to a contradiction. A similar contradiction is obtained if it is supposed that $t - (X + Y)$ lies without the simple closed curve $AXBYA$. Thus it follows that every component of K is either a point or a simple continuous arc t such that no point of t, save possibly its end-points, is a limit point of $K - t$. Therefore, by a theorem due to R. L. Moore and J. R. Kline,[7] there exists a simple continuous arc which contains K.

Since every simple continuous arc in the plane is a subset of some simple closed curve, we have the following immediate corollary.

COROLLARY. *Under the hypothesis of theorem 3, there exists a simple closed curve which contains K.*

Definitions.—A continuous curve M is said to be cyclicly connected[8] provided every two points of M lie together on some simple closed curve in M. If the cyclicly connected continuous curve C is a subset of a continuous curve M, then C is called a maximal cyclic curve of M if and only if C is not a proper subset of any other cyclicly connected continuous curve which is a subset of M.

THEOREM 4. *Let K be an irreducible cutting of a bounded continuous curve M between two points A and B of M, let R_a and R_b denote the components of $M - K$ containing A and B, respectively, let N be the boundary of any complementary domain D of M, and suppose that K contains more than one point. Then* (1) *K lies wholly in some maximal cyclic curve C of M,* (2) *$K \cdot N$ is either vacuous or it lies on some simple closed curve in N;* (3) *if every point of $K \cdot N$ is accessible from each of the sets R_a and R_b, then $K \cdot N$ contains not more than two points; and* (4) *if K is a subset of N, then K consists of exactly two points.*

Proof.—(1) If X and Y are any two points of K, then since $R_a + X + Y$ and $R_b + X + Y$ are connected[9] subsets of M having in common just the points X and Y, it is clear that X and Y are not separated in M by any single point of M. But if two points of a continuous curve M do not lie together in the same maximal cyclic curve of M, there exists[8] a point of M which separates these two points in M. Thus it follows that there exists a maximal cyclic curve C of M which contains every point of K.

(2) Suppose $K \cdot N$ is not vacuous. By (1) K is a subset of some maximal cyclic curve C of M. Let G denote the complementary domain of C which contains D. By a theorem of the author's[8] the boundary J of G is a simple closed curve. Since K is a subset of C, it is clear that J must contain $K \cdot N$ and must itself be a subset of N.

(3) Every point of $K \cdot N$ is accessible from D. Hence every point of $K \cdot N$ is accessible from each of the three mutually exclusive connected sets R_a, R_b and D, no one of which contains a point of $K \cdot N$. Therefore, by theorem 1, $K \cdot N$ contains not more than two points.

It is to be noted that, in consideration of the proof of theorem 1 and the fact that every point of K is a limit point of each of the sets R_a and R_b, the conclusion of (3) holds if it is assumed merely that there exists one of the sets R_a and R_b such that every point of $K \cdot N$ is accessible from this set.

(4) If K is a subset of N, then by (2), K lies on some simple closed curve in N. Hence K has no continuum of condensation, and by a theorem of R. L. Wilder's,[10] every point of K is accessible from each of the sets R_a and R_b. Therefore, by (3), K contains not more than two points; and since by hypothesis K contains more than one point, it follows that K consists of exactly two points. This completes the proof of theorem 4.

THEOREM 5. *Let the sets M, K, A, B, R_a, and R_b be defined exactly as in the statement of theorem 4. Suppose every point of K is accessible from each of the sets R_a and R_b. Then either K is a simple closed curve or there exists a simple continuous arc t which contains K but which contains no point of M not belonging to K, i.e., $t \cdot M \equiv K$.*

Proof.—(1) Let us suppose that K is not a simple closed curve and that it contains more than one point, since the case where K contains only one point is trivial. Then by theorem 3 there exists some simple continuous arc which contains K. Hence, every component of K is either a point or an arc, and no component of K separates the plane. Furthermore, the components of K form an upper semi-continuous[11] collection of continua. Therefore, as established by R. L. Moore,[11] if G denotes the collection whose elements are the components of K and all the *points* in the plane not belonging to K, then the sum of all the elements of G constitutes a space S' which is homeomorphic with the ordinary Euclidean point plane.

Then since K is a closed totally disconnected set of elements of G, and since, by hypothesis, $M - K$ is not connected, it follows by a theorem of

R. G. Lubben's[12] that there exists a simple closed curve J of elements of G which contains K but contains no point of $M - K$, and which encloses one of the sets R_a and R_b, say R_a, but does not enclose the other.

Now let H denote the *point set* consisting of all points X such that X belongs to some element of J. It is clear that H is a continuum. It must be a continuous curve. For if not, then[13] it contains a countable infinity of mutually exclusive continua $T, H_1, H_2, H_3, \ldots$, all of diameter $>$ some positive number given in advance, and such that T is the sequential limiting set of the sequence of continua $H_1, H_2, \ldots$. Now since K is closed, $H-K$ is the sum of a countable number of mutually exclusive segments $S_1, S_2, \ldots$ of J. Since no one of these segments contains a point of K, then for each i, the elements of S_i are all points, and hence S_i is a point set. Furthermore S_i belongs to some complementary domain D_i of M, and since, by theorem 4, part (3), the boundary N_i of D_i contains not more than two points of K, it follows that there exist two points A_i and B_i of K on N_i such that $S_i + A_i + B_i$ is closed and, in fact, is a simple continuous arc from A_i to B_i. Therefore, the sets $S_1, S_2, \ldots$ are ordinary arc segments. It readily follows that every point of T must belong to K, and since every component of K is either a point or an arc, T must be an arc. And since, by theorem 3, no interior point of T can be a limit point of $K-T$, it follows that there exist a subarc T^* of T and an infinite sequence $H_1^*, H_2^*, \ldots$ of continua all of diameter $>$ some positive number d given in advance, and such that, for every integer $i > 0$, H_i^* is a subset of H_i and contains no point whatever of K. Hence, for each i, the continuum H_i^* must belong to one of the segments $S_1, S_2, \ldots$. Since the segments $S_1, S_2, \ldots$ are ordinary arc segments, and since each of the continua $H_1^*, H_2^*, \ldots$ is of diameter $> d$, then no one of the segments $[S_i]$ can contain an infinite number of the continua $[H_i^*]$. It follows that infinitely many of the segments $[S_i]$ must be of diameter $> d$. But each of these segments lies in some complementary domain of M, no two of them can lie together in a single complementary domain of M, and only a finite number of the complementary domains of M are of diameter $> d$. Hence, not more than a finite number of the segments $[S_i]$ can be of diameter $> d$. Thus the supposition that H is not a continuous curve leads to a contradiction.

Now let I and E, respectively, denote the interior and exterior of J. Since neither I nor E contains a point of K, then the elements of I and of E are all points, and hence I and E are ordinary domains. And since I contains R_a and E contains R_b, it follows[9] that the point set H is the common boundary of the two domains I and E. But it was shown above that H is a continuous curve. Hence, by a theorem of R. L. Moore's,[14] H is a simple closed curve. Now let A_1 and B_1 denote the end-points of S_1, and let t denote that arc of H from A_1 to B_1 which does not contain S_1.

Then the arc t contains K but contains no point of $M-K$. This completes the proof of the theorem.

Note.—The above proof for theorem 5 is by no means an elementary one. The following proof, obtained by Mrs. G. T. Whyburn, is of interest both because of its more elementary character and because of the interesting way in which it utilizes theorem 3.

Proof.—(2) Let us suppose K is not a simple closed curve, and that it contains more than one point. By theorem 4, part (3), the boundary of no complementary domain of M contains more than two points of K. Let $D_1, D_2, \ldots$ be those complementary domains of M such that, for each integer $i > 0$, the boundary of D_i contains two distinct points A_i and B_i of K. For each $i > 0$, the domain D_i contains a segment S_i of an arc whose end-points are A_i and B_i.

Let N denote the set of points $R_a + R_b + K$. Clearly[9] N is a continuum. And since, by theorem 3, every component of K is either a point or an arc t no interior point of which is a limit point of $K - t$, it is readily shown that N must be a continuous curve. Now let $I_1, I_2, I_3, \ldots$ be the complementary domains of N such that, for each integer $i > 0$, the boundary of I_i contains two distinct points X_i and Y_i of K. Then there must exist at least one of the segments $[S_i]$ having X_i and Y_i as its end-points. For by theorem 4, part (2), X_i and Y_i lie together on some simple closed curve J_i in the boundary N_i of I_i. And of the two arcs t_1 and t_2 of J_i from X_i to Y_i, one of them must lie, save for the points X_i and Y_i, wholly in R_a and the other, save for its end-points, wholly in R_b; for if, on the contrary, one of the sets R_a and R_b, say R_a, contains points P_1 and P_2 belonging to t_1 and t_2, respectively, then $t_1 - (X_i + Y_i)$ and $t_2 - (X_i + Y_i)$ must each belong to R_a. Hence R_b must lie wholly either within or without J_i, say without J_i; then I_i must lie within J_i and if P_1P_2 denotes an arc[14] in R_a from P_1 to P_2, then P_1P_2 separates the exterior of J_i, and hence not both of the points X_i and Y_i can be accessible from R_b, contrary to hypothesis. Thus it follows that $t_1 - (X + Y)$ belongs to one of the sets R_a and R_b, say to R_a, and $t_2 - (X_i + Y_i)$ belongs to R_b. Then there must exist a complementary domain of M whose boundary contains both X_i and Y_i; for if not then M would contain[15] a simple closed curve C which separates X_i from Y_i, and clearly this is not possible, for C would contain an arc joining a point P of t_1 and a point Q of t_2 and lying otherwise wholly in I_i, which is absurd since P belongs to R_a and Q to R_b. Hence there exists a complementary domain D of M whose boundary contains both X_i and Y_i; accordingly there exists at least one of the segments $[S_i]$ whose end-points are X_i and Y_i. For each positive integer i, let us select just one segment S_i^* of the collection $[S_i]$ whose end-points are X_i and Y_i. Let E denote the sum of all such segments $[S_i^*]$ thus selected, and let T denote the point set $K + E$.

Since K is closed and bounded, and since not more than a finite number of the segments $[S_i^*]$ are of diameter $>$ any preassigned positive number, it readily follows that T is closed and bounded. And since T contains an arc segment S_i^* which joins every two points X_i and Y_i of K which lie together on the boundary of a complementary domain of N and which lies wholly in some complementary domain of M, it follows with the help of theorems by C. M. Cleveland[15] and R. L. Moore[16] that T is connected. Hence, T is a bounded continuum.

Now there does not exist a complementary domain D of the continuum $N + T$ which has a boundary point P_1 in R_a and also a boundary point P_2 in R_b. For suppose there does exist such a domain D. Then D lies wholly in some complementary domain G of N. (1) In case D is identical with G, then G can contain no point of E, and therefore the boundary U of G can contain not more than one point P of K. Now P must separate P_1 and P_2 in U; this being so, it readily follows that P also separates P_1 and P_2 in M. But since K, by hypothesis, is an irreducible cutting of M between A and B, then K must be an irreducible cutting of M between P_1 and P_2. Hence K must consist of a single point P. But we have supposed that K contains more than one point. Hence, in this case, the supposition that the domain D exists leads to a contradiction. (2) In case D is not identical with G, then G must contain one segment S^* of the collection $[S_i^*]$, and therefore the boundary F of G must contain the end-points X^* and Y^* of S^* which belong to K. Then since the sum of the points X^* and Y^* separates P_1 and P_2 in U, it readily follows that the segment S^* separates G into two domains the boundary of one of which contains P_1 but not P_2 and the boundary of the other contains P_2 but not P_1. But the boundary of D contains both P_1 and P_2. Hence, in any case, the supposition that the domain D exists leads to a contradiction.

Let G_a denote the sum of all those complementary domains of $N + T$ whose boundary contains a point of R_a, and G_b the sum of all those whose boundary contains a point of R_b. Then from what we have just shown it follows that $R_a + G_a$ and $R_b + G_b$ are mutually exclusive connected point sets. Now if X is any point of T which belongs to K, then X is accessible from each of the sets $R_a + G_a$ and $R_b + G_b$, for by hypothesis it is accessible from R_a and also from R_b and clearly some point of R_a (R_b) is accessible from each component of G_a (G_b). And if P is any point of T not belonging to K, then P belongs to some segment S^* of $[S_i^*]$. The segment S^* separates some complementary domain G of N into two domains G_1 and G_2 one of which, it is clear from the second paragraph of this proof, must belong to G_a and the other to G_b. Hence, P is accessible from each of the sets $R_a + G_a$ and $R_b + G_b$. Therefore, since T is connected, it follows by theorem 3 that T must be either an arc or a simple closed curve. If T is a simple closed curve, then since K is not a simple closed curve,

there exists a component S of $T-K$. And $T-S$ is, then, a simple continuous arc which contains K but no point of $M-K$. This completes the proof.

In proof (1) of theorem 5 we have established the following theorem.

THEOREM 6. *Under the hypothesis of theorem 5, there exists a simple closed curve J which contains K but no point of $M-K$ and which separates R_a and R_b.*

THEOREM 7. *If, in addition to the hypothesis in theorem 5, it is assumed that the boundaries of no two of the complementary domains of M have a common point, then there exists a simple closed curve which contains K and is a subset of M.*

Theorem 7 is readily established with the aid of theorem 6.

THEOREM 8.—*Let the sets M, A, B, K, R_a and R_b be defined exactly as in the statement of theorem 4. Then if every component T of K is a continuous curve and there exist two non-cut points of T each of which is accessible from both of the sets R_a and R_b, there exists a simple closed curve which contains K but no point of $M-K$ and which separates R_a and R_b.*

Proof.—It follows by a theorem of W. L. Ayres[17] that either K is itself a simple closed curve or each component T of K is either a point or an arc. If K is itself a simple closed curve, our theorem is an immediate consequence of theorem 1. So let us suppose K is not a simple closed curve. Then if T is any component of K which consists of more than one point, T must be an arc whose end-points are accessible from each of the sets R_a and R_b. Then by almost identically the same argument as used in the third paragraph of the proof of theorem 3 it is shown that no interior point of T is a limit point of $K-T$. Hence it follows that K can have no continuum of condensation; and since K is the M-boundary of each of the sets R_a and R_b, then by a theorem of R. L. Wilder's[10] *every point* of K must be accessible from each of the sets R_a and R_b. Therefore, by theorem 6, there exists a simple closed curve which contains K but no point of $M-K$ and which separates R_a and R_b.

In conclusion I will point out that if M is a continuous curve every subcontinuum of which is a continuous curve, then[5] the accessibility hypotheses in theorems 4, 5, 6 and 8 are always satisfied.

* The Editor regrets to state that this article was lost in transmission from author to editorial office—which accounts for the delay in publication.—E. B. W.

[1] Presented to the American Mathematical Society, December 28, 1927.

[2] Schoenflies, A., *Gottingen Nachrichten*, **1902**, p. 185. Cf. also, Moore, R. L., *Trans. Amer. Math. Soc.*, **17**, 158, 159 (1916); Swingle, P. M., *Bull. Amer. Math. Soc.*, **32**, 110–11 (1926) (abstract), and a remark by C. Kuratowski in *Fund. Math.*, **6**, 143 (1924).

[3] Cf. Whyburn, G. T., *Bull. Amer. Math. Soc.*, **33**, 408 (1927) (abstract). Paper to appear in full in *Fund. Math.*, **13**.

[4] A component of a point set K is a maximal connected subset of K; cf. Hausdorff, *Mengenlehre*, **1927**.

[5] Cf. Whyburn, G. T., "Concerning the Complementary Domains of Continua," forthcoming in the *Annals of Mathematics*, theorem 15. For the theorem for the case where the sets R_1, R_2 and R_3 are open subsets of M, see Whyburn, G. T., these PROCEEDINGS, **13**, 650–657 (1927), theorem 3.

[6] Moore, R. L., *Trans. Amer. Math. Soc.*, **21**, 333–347 (1920).

[7] Moore, R. L., and Kline, J. R., *Annals of Mathematics*, **20**, 182–223 (1919).

[8] Cf. Whyburn, G. T., these PROCEEDINGS, **13**, 31–38 (1927) and *Bull. Amer. Math. Soc.*, **33**, 305–308 (1927).

[9] This follows from the fact that K is the M-boundary of each of the sets R_a and R_b. Cf. the abstract of my paper "On Irreducible Cuttings of Continua," presented to the American Mathematical Society, December 28, 1927.

[10] Wilder, R. L., *Fund. Math.*, **7**, 340–377 (1925), theorem 1.

[11] Moore, R. L., *Trans. Amer. Math. Soc.*, **27**, 416–428 (1925).

[12] Lubben, R. G., *Bull. Amer. Math. Soc.*, **32**, 114 (1926) (abstract).

[13] Moore, R. L., *Bull. Amer. Math. Soc.*, **29**, 291–297 (1923).

[14] Moore, R. L., *Math. Zeit.*, **15**, 254–260 (1922).

[15] Cleveland, C. M., these PROCEEDINGS, **13**, 275–276 (1927).

[16] Moore, R. L., *Ibid.*, **13**, 711–716 (1927), theorem 2.

[17] Ayres, W. L., *Bull. Amer. Math. Soc.*, **33**, 565–571 (1927), theorems 4 and 5. For an additional hypothesis needed in these theorems, see a note in the *Bulletin of the American Mathematical Society*, **34**, 107–108 by W. L. Ayres. It is of interest to note that while, as shown by theorem 8 of the present paper, Ayres' condition (4) in his note makes the point set K such that it is a subset of a simple closed curve, his condition (4′) does not do this. His condition (4) implies (4′) but (4′) does not imply (4).

INCOMPLETE SYSTEMS OF PARTIAL DIFFERENTIAL EQUATIONS[1]

BY JOSEPH MILLER THOMAS

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA

Communicated July 5, 1928

The theory of a system of total differential equations

$$\frac{\partial u_\alpha}{\partial x_i} = a_{\alpha i} \qquad \alpha = 1, 2, \ldots, r;\ i = 1, 2, \ldots, n \tag{1}$$

which are completely integrable, that is, for which the conditions

$$\frac{\partial a_{\alpha i}}{\partial x_j} + \sum_{\beta=1}^{r} \frac{\partial a_{\alpha i}}{\partial u_\beta} a_{\beta j} = \frac{\partial a_{\alpha j}}{\partial x_i} + \sum_{\beta=1}^{r} \frac{\partial a_{\alpha j}}{\partial u_\beta} a_{\beta i} \tag{2}$$

are satisfied identically in u and x, is known to correspond exactly to the theory of a jacobian system of simultaneous linear homogeneous equations,[2] but the usual treatment of the incompletely integrable case for the two

types of system does not preserve this correspondence. It is the purpose of the present paper to give a treatment displaying the equivalence of the two sorts of systems in the incompletely integrable case.

Consider a simultaneous system of n independent linear homogeneous equations of the first order in a single unknown function f. Without loss of generality we can assume the system in the form

$$X_i f \equiv \frac{\partial f}{\partial x_i} + \sum_{\alpha=1}^{r} a_{\alpha i} \frac{\partial f}{\partial u_\alpha} = 0. \qquad i = 1, 2, \ldots, n. \tag{3}$$

Any solution of (3) will also satisfy

$$(X_i X_j) f \equiv \sum_{\beta=1}^{r} b_{\beta ij} \frac{\partial f}{\partial u_\beta} = 0, \tag{4}$$

where the parentheses denote the Poisson operator. Equations (4) contain none of the derivatives $\partial f/\partial x_i$ and, therefore, cannot be algebraic consequences of (3). Either they are identically satisfied, or the system is incomplete. In the former case, the coefficients of $\partial f/\partial u$ in (4) vanish, that is, the relations

$$\frac{\partial a_{\alpha i}}{\partial x_j} + \sum_{\beta=1}^{r} a_{\beta j} \frac{\partial a_{\alpha i}}{\partial u_\beta} = \frac{\partial a_{\alpha j}}{\partial x_i} + \sum_{\beta=1}^{r} a_{\beta i} \frac{\partial a_{\alpha j}}{\partial u_\beta} \tag{5}$$

are satisfied identically in u and x. System (3) is then jacobian and has r integrals distinct with respect to the variables $u_1, u_2, \ldots, u_r$. In the latter case, equations (5) constitute additional conditions on f, and (3) has fewer than r distinct integrals of the type ordinarily considered, namely, functions of u and x each of which substituted in (3) for f reduces the left members to zero identically in u and x.

We may seek, however, an integral of (3) in an extended sense, that is, a function f and a set of relations among the variables u and x having the following properties: the relations are distinct with respect to the u's; and when f is substituted in (3), equations (3) are satisfied, not necessarily identically, but as a consequence of the relations among u and x. Any such integral must satisfy (4) in the same manner.

If we require that (3) have r integrals of this type distinct with respect to the u's, then regarding i and j as temporarily fixed in the linear homogeneous equations

$$\sum_{\beta=1}^{r} b_{\beta ij} \frac{\partial f_\alpha}{\partial u_\beta} = 0, \qquad \alpha = 1, 2, \ldots, r,$$

we conclude from the non-vanishing of $|\partial f/\partial u|$ that $b_{\beta ij} = 0$, or that (5)

must be satisfied. We denote this system by $F_1 = 0$, and build up a sequence of relations

$$F_1 = 0, \quad F_2 = 0, \ \ldots, \quad F_N = 0, \tag{6}$$

in which the set F_α comprises all the equations $X_i(F_{\alpha-1}) = 0$. Since $X_i(0) = 0$, our integrals must satisfy (6).

There must exist an integer N for which the corresponding systems (6) are equivalent to a system

$$f_\beta(u, x) = 0, \qquad \beta = k+1, \ldots, r \tag{7}$$

which can be solved for $r-k$ of the u's, say $u_{k+1}, \ldots, u_r$, in terms of the remaining u's and the x's and whose solution substituted in the equations of $F_{N+1} = 0$ converts them into identities in x and $u_1, u_2, \ldots, u_k$. To prove this a necessary condition for the existence of the integrals of (3), let us choose any value for N, and let the functional matrix of the corresponding system (6) with respect to the u's have the rank l. We can solve certain of equations (6) for l of the u's, say $u_{r-l+1}, \ldots, u_r$. The results, on being substituted in *any* of equations (6), give identities in $u_1, \ldots, u_{r-l}$, and hence must also give identities in x because no relation among the x's is implied by the set of relations among u and x admitted. The expressions for $u_{r-l+1}, \ldots, u_r$ when substituted in $F_{N+1} = 0$ must give *either* identities in $u_1, \ldots, u_{r-l}$ and x *or* relations involving $u_1, \ldots, u_{r-l}$ since relations among the x's alone are excluded. If nothing but identities result from the substitution, the N in question has the properties announced at the beginning of this paragraph, and $k = r-l$; if some relations in $u_1, u_2, \ldots, u_{r-l}$ result, then these new relations treated as (6) were will determine certain of the remaining u's. We substitute them together with the u's previously obtained in $F_{N+2} = 0$. This process must ultimately lead to nothing but identities because there is only a finite number of u's.

Conversely, suppose that this necessary condition is fulfilled. By hypothesis, any solution of (7) satisfies $X_i F_\beta = 0$, that is,

$$\frac{\partial f_\beta}{\partial x_i} + \sum_{\gamma=1}^{k} a_{\gamma i} \frac{\partial f_\beta}{\partial u_\gamma} + \sum_{\gamma=k+1}^{r} a_{\gamma i} \frac{\partial f_\beta}{\partial u_\gamma} = 0. \tag{8}$$

Differentiating (7) and substituting $\partial f_\beta/\partial x$ and $\partial f_\beta/\partial u_\gamma$, $\gamma = 1, 2, \ldots, k$, in (8) we get

$$\sum_{\gamma=k+1}^{r} \frac{\partial f_\beta}{\partial u_\gamma} \left(a_{\gamma i} - \frac{\partial u_\gamma}{\partial x_i} - \sum_{\rho=1}^{k} a_{\rho i} \frac{\partial u_\gamma}{\partial u_\rho} \right) = 0, \quad \beta = k+1, \ldots, r$$

whence

$$a_{\gamma i} = \frac{\partial u_\gamma}{\partial x_i} + \sum_{\rho=1}^{k} a_{\rho i} \frac{\partial u_\gamma}{\partial u_\rho}. \qquad \gamma = k+1, \ldots, r. \tag{9}$$

Suppose the values of $u_{k+1}, \ldots, u_r$ substituted in the coefficients of (3) and seek a function f of $x, u_1, u_2, \ldots, u_k$ alone satisfying the resulting equations. The equations defining f will be

$$\frac{\partial f}{\partial x_i} + \sum_{\rho=1}^{k} A_{\rho i} \frac{\partial f}{\partial u_\rho} = 0, \tag{10}$$

where the A's represent what the a's become after the substitution. The conditions that (10) be jacobian are

$$\frac{\partial A_{\rho i}}{\partial x_j} + \sum_{\lambda=1}^{k} A_{\lambda j} \frac{\partial A_{\rho i}}{\partial u_\lambda} = \frac{\partial A_{\rho j}}{\partial x_i} + \sum_{\lambda=1}^{k} A_{\lambda i} \frac{\partial A_{\rho j}}{\partial u_\lambda}. \tag{11}$$

Now we have

$$\frac{\partial A_{\rho i}}{\partial x_j} = \frac{\partial a_{\rho i}}{\partial x_j} + \sum_{\beta=k+1}^{r} \frac{\partial a_{\rho i}}{\partial u_\beta} \frac{\partial u_\beta}{\partial x_j}, \qquad \frac{\partial A_{\rho i}}{\partial u_\lambda} = \frac{\partial a_{\rho i}}{\partial u_\lambda} + \sum_{\beta=k+1}^{r} \frac{\partial a_{\rho i}}{\partial u_\beta} \frac{\partial u_\beta}{\partial u_\lambda},$$

it being understood that $u_{k+1}, \ldots, u_r$ have been replaced by their values from (7) in the right members. The left member of (11) is therefore

$$\frac{\partial a_{\rho i}}{\partial x_j} + \sum_{\lambda=1}^{k} a_{\lambda j} \frac{\partial a_{\rho i}}{\partial u_\lambda} + \sum_{\beta=k+1}^{r} \left(\frac{\partial u_\beta}{\partial x_j} + \sum_{\lambda=1}^{k} a_{\lambda j} \frac{\partial u_\beta}{\partial u_\lambda} \right) \frac{\partial a_{\rho i}}{\partial u_\beta}.$$

By use of (9), this becomes

$$\frac{\partial a_{\rho i}}{\partial x_j} + \sum_{\beta=1}^{r} a_{\beta j} \frac{\partial a_{\rho i}}{\partial u_\beta}.$$

Consequently (11) are simply certain ones of (5) with $u_{k+1}, \ldots, u_r$ eliminated by means of (7). But by hypothesis such elimination leads to identities in $x, u_1, \ldots, u_k$. Hence, system (10) is jacobian and admits k integrals $f_1, f_2, \ldots, f_k$ distinct with respect to $u_1, u_2, \ldots, u_k$. These taken together with (7) will be called a *complete set of integrals*. It is to be noted that this set reduces to the ordinary complete set if the system (3) is jacobian.

We have, therefore, proved

THEOREM. *A necessary and sufficient condition for the existence of r independent integrals (in the extended sense) of (3) is the existence of an N such that equations (6) are consistent for the determination of some or all of the u's and such that their solution satisfies $F_{N+1} = 0$.*

$f_1 - c_1$ is, of course, an integral of (10), c_1 being an arbitrary constant. Put

$$f_1 - c_1 = 0, \quad f_2 - c_2 = 0, \ldots, \quad f_k - c_k = 0; \tag{12}$$

these equations and (7) can be solved to give $u_1, u_2, \ldots, u_r$ as functions of x and the arbitrary constants c.

To apply these results to system (1), suppose u's satisfying (1) are defined by a set of relations

$$f_\alpha (u, x) = 0. \qquad \alpha = 1, 2, \ldots, r. \tag{13}$$

There results

$$\frac{\partial f_\alpha}{\partial x_i} + \sum_{\beta=1}^{r} a_{\beta i} \frac{\partial f_\alpha}{\partial u_\beta} = 0, \tag{14}$$

that is, the $f_1, f_2, \ldots, f_r$ constitute a set of integrals of (3) of the type just considered. Consequently, to every system (1) there is an equivalent system (3), and vice versa; and the general solution of (1) is given by (7) and (12).

The integrability conditions (2) and (5) are identical. If a system of type (3) is jacobian, the corresponding system (1) is completely integrable. Therefore, if the relations like (6) be employed to eliminate $u_{k+1}, \ldots, u_r$ from those equations of system (1) which contain derivatives of $u_1, u_2, \ldots, u_k$, the resulting system in $u_1, u_2, \ldots, u_k$ is completely integrable, and any solution of it taken with $u_{k+1}, \ldots, u_r$ as determined by the relations like (6) satisfies all the equations of (1).

If we note that applying the operator X_i to a function of u and x is the same as differentiating that function with respect to x_i and eliminating first derivatives by means of (1), we see that the theorem given above is equivalent to a known theorem[3] for systems (1).

We therefore have developed a theory of systems (3) which corresponds to the theory of systems (1) in the incompletely integrable case. It would seem that no extension of the theory of systems (1) is possible to correspond to the treatment of system (3) by adjoining to it (4) and finally arriving at an equivalent jacobian system: for in the treatment of (1) we are only interested in satisfying (14) by virtue of (13) and not in satisfying (14) identically.

[1] The first draft of this paper was written while the author was National Research Fellow at Princeton University.

[2] See Goursat, *Leçons sur l'intégration des équations aux dérivées partielles du premier ordre*, Paris, 1921, p. 103.

[3] O. Veblen and J. M. Thomas, *Annals Math.*, **27**, 1926 (288).

THE EFFECT OF TEMPERATURE ON THE VIABILITY OF SUPERFEMALES IN DROSOPHILA MELANOGASTER

BY TH. DOBZHANSKY*

COLUMBIA UNIVERSITY

Communicated June 20, 1928

The superfemale is an individual having three X-chromosomes and two sets of autosomes. It has also many external characteristics by means of which it can be easily distinguished both from the normal sexes and from other sexual forms, i.e., intersexes and supermales. The superfemales arise in several kinds of genetical experiments concerned with abnormal chromosomal situations, such as in the progeny of the attached X-chromosome female, of the triploid female and in cases of non-disjunction of the X-chromosomes. But the frequency of superfemales found in cultures is always very much below that which might be expected theoretically on the basis of a knowledge of the genetic situation in a given case. Therefore, the superfemale can be considered as a typical semilethal form, which only relatively rarely reaches the imago stage.

Another fact, which can easily be noticed in working with the cultures producing superfemales, is the considerable variability of their frequency in the progeny of flies of like genetical constitution. This variability may be explained by the effect of some external influences on the viability of superfemale individuals. The preliminary results of Dr. A. H. Sturtevant (unpublished) suggested that a change of temperature really affects the frequency of the superfemales. The purpose of the present investigation is to check up this point more definitively.

The easiest way to obtain superfemales is to breed attached-X females (described by L. V. Morgan, 1922; T. H. Morgan, Bridges, Sturtevant, 1925). Such a female contains in its cells two X-chromosomes attached by their ends to each other, and in addition one Y-chromosome and two sets of the autosomes. In maturation of gametes two kinds of eggs are formed; one containing two attached X-chromosomes, and the other only the Y-chromosome, both having one set of the autosomes. Fertilized by the X-containing sperm of a normal male these eggs give two kinds of zygotes: XXX, which is the superfemale, and XY, which is normal male. In a similar way the Y-containing sperm gives XXY female, which is like the mother, and YY zygotes, which presumably never survive. If, as in the strain used by the author, the mother has the genes for the yellow recessive body color in both of its attached X-chromosomes, and the father is wild type with respect to sex-linked genes, all the daughters will be yellow and all the sons wild type in appearance. The superfemales in this

case must also be wild type because of the dominance of one wild type over two yellow allelomorphs.

Females from the attached-X stock were crossed to sooty males in order to prevent the possible harmful effect of the inbreeding on the viability of the flies. Sooty is a recessive autosomal gene and its presence in the culture does not affect the accuracy of classification of the sexes or of the yellow body color. The progeny of this cross was bred in pairs and placed in common half-pint bottles provided with a standard amount of the cornmeal-molasses food. The flies were raised at four different temperatures: $30^1/_2°$, $24^1/_2°$, 20° and 16°C. The accuracy of the temperature control was $\pm^1/_4°$ at $30^1/_2°$ and 20°, and $\pm^1/_2°$ in the other two series. The flies were allowed to lay eggs until pupae of the next generation appeared, i.e., over a longer period in low and a shorter one in high temperature. The hatching flies were counted once every two days until the end of hatching. The summary of all counts is given in table 1. The percentage of superfemales is shown in table 2.

TABLE 1

THE PROGENY OF ATTACHED-X FEMALES AT DIFFERENT TEMPERATURES

			DETACHMENT					
T°	YELLOW FEMALES	WILD TYPE MALES	SUPER-FEMALES	WILD TYPE FEMALES	YELLOW MALES	EDUCATIONAL EXCEPTIONS IN THE MALE	INTER-SEXES	TOTAL
$30^1/_2$	1749	1878	1	5	1	.	1	3635
$24^1/_2$	3262	3645	101	2	1	.	.	7011
20	3013	3340	351	1	.	.	.	6705
16	2635	2630	8	5	3	2	.	5283
Total	10659	11493	461	13	5	2	1	22634

TABLE 2

THE PERCENTAGE OF SUPERFEMALES AND THE SEX RATIO AT DIFFERENT TEMPERATURES

T°	SUPERFEMALES	FEMALE:MALE RATIO	DETACHMENT, %
$30^1/_2$	0.03 ± 0.03	1:1.07	0.17
$24^1/_2$	1.44 ± 0.14	1:1.12	0.04
20	5.23 ± 0.27	1:1.11	0.01
16	0.15 ± 0.05	1:1.00	0.15

In the cultures raised at $30^1/_2°$ there were found only $0.03 \pm 0.03\%$ of superfemales. This means practically a complete non-occurrence of this kind of individuals. At $24^1/_2°$ the superfemales were not rare ($1.44 \pm 0.14\%$) and usually could be found in each culture bottle. At 20° the highest frequency of the superfemales was observed ($5.23 \pm 0.27\%$). The differences between these three figures are indubitably significant from the statistical standpoint. Below 20° the frequency of the superfemales diminishes again and at 16° only $0.15 \pm 0.05\%$ of individuals found in the cultures are superfemales. This percentage is nearly as low as that observed at $30^1/_2°$ and the difference between these two values is below the limit of certainty.

The results obtained may be illustrated in a different way. Assuming that the attached-X female produces XX and Y gametes in equal numbers, and that all the XXX zygotes may survive to the imago stage, a frequency 33.3% of the superfemales may be expected in the cultures. The frequencies actually observed are only fractions of that. At $30^1/_2°$ only 1 of 1111.1 of the superfemales survives. For the temperatures $24^1/_2°$, 20° and 16° this ratio is 1:23.1, 1:6.4 and 1:222.2, respectively.

Two hypotheses can be advanced for the explanation of these facts. First, the temperature may affect the frequency of production of XX and Y gametes in the attached-X mother. If more XX are produced at 20° and at $30^1/_2°$ and 16° more Y gametes, then the superfemales must be rarer in the last two temperatures and more frequent at 20°. But in such a case the sex ratio must also be affected. At 20° must be found a very considerable excess of females, and at $30^1/_2°$ and at 16° an excess of males. A disturbance of the sex ratio due to production of different numbers of XX and Y gametes in the attached-X female was described by T. H. Morgan, (Morgan, Sturtevant, Bridges, 1927). As table 2 shows, in my case the sex ratio is nearly equal at all the temperatures used, or even slightly changed toward the male side at 20° and $24^1/_2°$. Since no correlation is observed between the sex ratio and the frequency of the superfemales also in the separate cultures, this hypothesis is untenable.

TABLE 3

THE NUMBER OF OFFSPRING OF ONE ATTACHED-X FEMALE AT DIFFERENT TEMPERATURES

T°	MEAN VALUE	σ	LIMITS
$30^1/_2$	86.6±6.2	40.2	18–175
$24^1/_2$	188.8±10.5	33.8	9–295
20	203.5±4.1	23.3	151–241
16	159.8±6.1	34.8	27–217

Second, the temperature may affect the viability of the superfemales themselves. The XXX zygotes may be absolutely lethal at $30^1/_2°$ and at 16°, but some of them may survive at 20 or $24^1/_2°$. There is some independent evidence in favor of this hypothesis. The superfemales hatched at $24^1/_2°$ and at 16° are extremely weak, they very often have misshaped legs, shriveled wings and defects in the arrangement of the abdominal tergites. Some individuals break their pupa-cases but can not get out and die half emerged. The superfemales hatched at 20° are stronger and the defects just mentioned are seldom observed. Dead half-emerged pupae were never found at 20°.

Many dead pupae can be observed in the bottles kept at all the temperatures used, but at $30^1/_2°$ and at 16° they seem to be more frequent than at $24^1/_2°$ and especially at 20°. The fact that these dead pupae contained

superfemales was several times tested directly by means of dissection: inside of the pupa-cases fully formed dead flies were found. Li (1927) has shown also that approximately 25% of the eggs laid by an attached-X mother die without development (these are presumably the YY eggs), and another 25% die in old larval or pupal stages (superfemales).

To what extent the temperature affects the viability of the superfemales directly or indirectly cannot be decided at present. It is possible that the effect of temperature is partially due to changing of the physico-chemical condition of the food. These conditions may be the most favorable at 20° for all the flies and for the superfemales in particular. Comparing the average numbers of flies produced per bottle at different temperatures (Table 3) with the percentage of the superfemales (Table 1), it can be seen that there is some correlation between these values. But the average number of flies per bottle depends upon both egg-laying rate and laying duration. Therefore no conclusion can be drawn on the basis of this correlation. No correlation exists between the number of individuals hatched in a separate bottle and the percentage of the superfemales in the same bottle (at 20° the coefficient of correlation is equal to $+0.07 \pm 0.18$).

As seen from the table 1, some exceptional wild type females and yellow males are present in the cultures. The yellow males and most of the wild type females (giving some yellow sons if crossed to wild males) are the result of breaking apart of the attached pair of the X-chromosomes in the gametogenesis of their mother. Two wild type females have given no yellow males in their progeny when tested. Their origin is due probably to the equational non-disjunction of the X-chromosomes in the spermatogenesis of their father.

The percentage of the exceptional individuals due to the detachment of the attached X-chromosomes seems to be higher at both $30^1/_2°$ and 16°, than at $24^1/_2°$ and 20°. This means probably that the high and the low temperature increase the frequency of breakage of the attached-X pair. This fact in the case of high temperature was already stated by Bridges and Gabritschevsky (1928). But in my case the low temperature seems to produce the same effect as the high.

Summary.—The viability of the superfemales is much affected by the temperature; the optimum lies near 20°. At $30^1/_2°$ and at 16° no, or very few, superfemales hatch from the pupae. Both high and low temperature increase the breaking apart of the attached X-chromosomes.

* Fellow of the International Education Board.

Bridges, C. B., and E. Gabritschevsky, "The Giant Mutation in Drosophila Melanogaster. I. The Heredity of the Giant," *Zeitschr. indukt. Abstamm. Vererb.*, **46**, 231–47 (1928).

Li, J. C., "The Effect of Chromosomal Aberrations on Development in Drosophila Melanogaster," *Genetics*, **12**, 1–58 (1927).

Morgan, L. V., "Non-criss-cross Inheritance in Drosophila Melanogaster," *Biolog. Bull.*, **42**, 267–74 (1922).

Morgan, T. H., C. B. Bridges and A. H. Sturtevant, "The Genetics of Drosophila," *Bibliographica Genetica*, **2**, 1–262 (1925).

Morgan, T. H., A. H. Sturtevant and C. B. Bridges, "The Construction of the Germ Material in Relation to Heredity," *Carnegie Inst. Year Book*, **26**, 284–8 (1927).

PRELIMINARY NOTE ON THE ELASTIC HYSTERESIS OF THE HUMAN AORTA

BY A. V. BOCK, P. S. BAUER AND J. H. MEANS

MEDICAL LABORATORY OF THE MASSACHUSETTS GENERAL HOSPITAL AND CRUFT LABORATORY, HARVARD UNIVERSITY

Communicated July 13, 1928

Introduction.—As a result of a study of blood pressure changes occurring during muscular exercise[1] an attempt has been made to determine the rôle played by the arch of the aorta under conditions producing an increase in

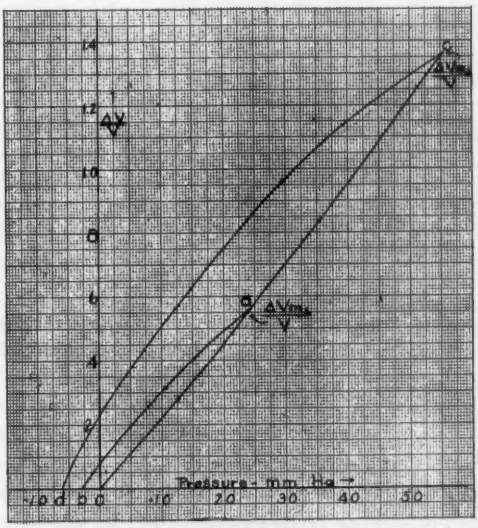


FIGURE 1

Typical curve illustrating elastic hysteresis of human aorta.

systemic blood pressure. Experimental results with reference to changes in blood pressure during exercise in different subjects suggested the possibility that variations in the elastic properties of the arch of the aorta

might be chiefly responsible for the reactions observed. Previous work[2] done on this subject has not related functional changes in the aorta to the mechanism of blood flow and blood pressure changes and to the limitations that changes in the arch of the aorta impose upon the body as a whole. The present communication is concerned only with certain aspects of the phenomena of hysteresis in the arch of the aorta.

Apparatus and Procedure.—The arch of the aorta (from aortic valve to first intercostal arteries) under examination is tied on to a water-mercury manometer and, after removal of the air from the system, the internal and external pressure on the aorta is adjusted to equality. The internal pressure is then increased monotonically to a certain maximum value. The pressure is then decreased until the initial volume is attained. At each step, the increment of the internal volume of the aorta is determined from a volumetric scale on one of the branches of the manometer. In this manner curves are obtained of the fractional increment of internal volume against the internal pressure. The process is then repeated, going to a higher maximum pressure.

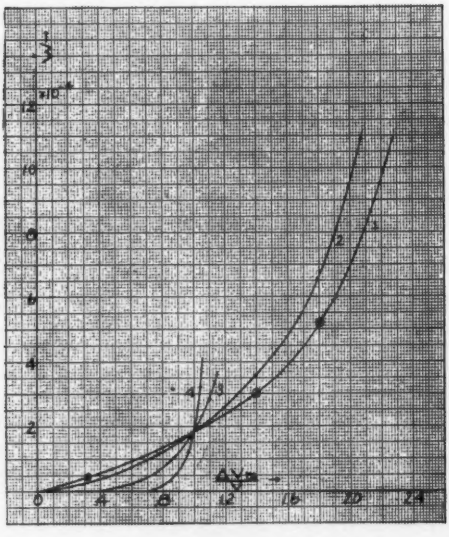


FIGURE 2

Curves of energy lost in heat in gram-calories per cycle per unit internal volume against maximum fractional increase of internal volume. Circles denote values computed from Equation 2. Dots represent experimental values.

Results.—Following the procedure outlined above, typical curves thus obtained are shown in figure 1. Curve *oab* was first obtained and then curve *ocd*. For the latter curve, the true origin is at the point *b*, but in order to preserve symmetry the curve has been translated to the right, placing the origin at the point *o*.

Now it can be shown that the area enclosed by the cyclic curve, *oab*, is proportional to the energy lost in heat by the stretching of the aorta to the maximum fractional increase of internal volume, $\frac{\Delta Vm}{V}$, where V is the unstretched internal volume.

The energy, W, lost in gram-calories per cycle per unit internal volume is given by:

$$W = k \times h \times A^1 \times 3185 \times 10^{-8}, \qquad (1)$$

where A^1 is the area enclosed in square inches, h and k are dimensional constants for the ordinates and abscissae.

If we now compute W for each maximum fractional increase of internal volume, $\frac{\Delta Vm}{V}$, we have as an empirical functional relationship, the typical curves illustrated in figure 2. Assuming that the relation between W and $\frac{\Delta Vm}{V}$ is analogous to the Steinmetz Law for magnetic hysteresis and is of the form:

$$W = A\left(\frac{\Delta Vm}{V}\right)^b, \qquad (2)$$

we obtain the computed values shown in the graph.

In every case we find that the constant A, which dimensionally is the energy lost in gram-calories per unit internal volume per cycle for a 100 per cent increase of internal volume, is the same. Descriptive data for the curves of figure 2 are given below:

CURVE	AGE	SEX	UNSTRETCHED INTERNAL VOLUME
1	15 months	male	2.0 cc.
2	41 years	male	14.0 cc.
3	78 years	male	15.4 cc.
4	65 years	female	21.2 cc.

The essential difference between aortae from the standpoint of hysteresis loss, lies in the magnitude of the exponent b.

In order to ascertain the effect of hysteresis loss on the work done by the heart, the efficiency of the aorta has been determined. We shall define efficiency as follows: Fractional efficiency is the ratio of the work delivered by the aorta on contraction from a certain maximum fractional increase in volume to the work done on the aorta in expanding it to the same value.

Typical curves of the efficiency against the maximum fractional increase in volume are shown in figure 3.

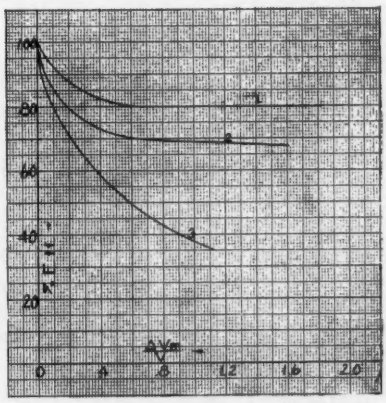


FIGURE 3
Typical cruves of efficiency against maximum fractional increase of internal value.

Let us consider curve 3 in figure 3 and reason qualitatively. We shall assume that the arterial part of the periphery obeys the same law as that for the arch of the aorta. Further we shall assume that the metabolic demand of the system in man at the age of 78 years remains the same as before advanced changes occurred in the arterial walls. Then with a 40 per cent increase of internal volume per systole, or a 20 per cent increase in average diameter, 45 per cent of the energy output of the heart is lost in heat compared with 25 per cent lost at the age of 41 years. The result is that the work of the heart must increase 20 per cent as a result of arteriosclerosis in order to maintain the same blood flow.

The subject will be treated more fully in a later communication.

[1] "Studies in Muscular Activity. III. Dynamical Changes Occurring in Man at Work," A. V. Bock, D. B. Dill, C. VanCaulaert, A. Fölling, and L. M. Hurxthal. (To be published.)

[2] (*a*) A. Fleisch, *Arch. f. d. ges. Physiol.*, **183**, 71 (1920). (*b*) R. Tigerstedt, *Die Physiologie des Kresslaufes*, Vol. 3, Section 106, pages 36–41. Published in Berlin and Leipsic, 1922.